PREVENCION Y TRATAMIENTO DE LA MALARIA • LA GUERRA DEL DESIERTO



INVESTIGACION y CIENCIA

Edición española de SCIENTIFIC AMERICAN

Febrero de 2006

Número 353















DIRECTOR GENERAL José M.ª Valderas Gallardo
DIRECTORA FINANCIERA Pilar Bronchal Garfella
EDICIONES Juan Pedro Campos Gómez
Laia Torres Casas
PRODUCCIÓN M.ª Cruz Iglesias Capón
Albert Marín Garau
SECRETARÍA Purificación Mayoral Martínez
ADMINISTRACIÓN Victoria Andrés Laiglesia
SUSCRIPCIONES Concepción Orenes Delgado
Olga Blanco Romero
EDITA Prensa Científica, S.A. Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 934 143 344 Telefax 934 145 413
www.investigacionyciencia.es

SCIENTIFIC AMERICAN

EDITOR IN CHIEF John Rennie
EXECUTIVE EDITOR Mariette DiChristina
MANAGING EDITOR Ricki L. Rusting
NEWS EDITOR Philip M. Yam
SPECIAL PROJECTS EDITOR Gary Stix
SENIOR EDITOR Michelle Press
SENIOR WRITER W. Wayt Gibbs
EDITORS Mark Alpert, Steven Ashley, Graham P. Collins, Steve Mirsky, George Musser y Christine Soares
PRODUCTION EDITOR Richard Hunt
GENERAL MANAGER Michael Florek
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL Dean Sanderson
PRESIDENT AND CHIEF EXECUTIVE OFFICER Gretchen G. Teichgraeber
CHAIRMAN John Sargent

DISTRIBUCION

para España:

LOGISTA, S. A.
Pol. Ind. Polvoranca
Trigo, 39, Edif. 2
28914 Leganés (Madrid)
Teléfono 914 819 800

para los restantes países:

Prensa Científica, S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona

PUBLICIDAD

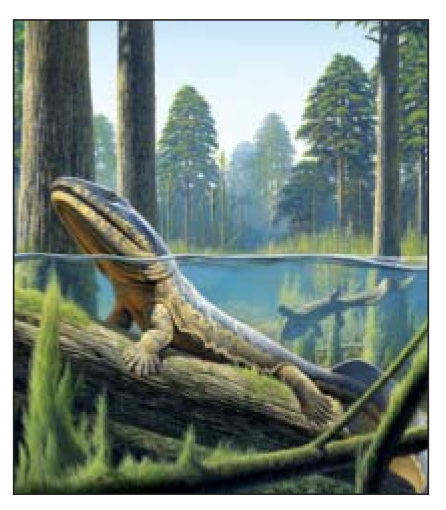
Madrid:
MOSAICO COMUNICACION, S. L.
Santiago Villanueva Navarro
Tel. y fax 918 151 624
Móvil 661 472 250
mosaicocomunicacion@yahoo.es

Cataluña:
QUERALTO COMUNICACION
Julián Queraltó
Sant Antoni M.ª Claret, 281 4.º 3.ª
08041 Barcelona
Tel. y fax 933 524 532
Móvil 629 555 703

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

M.ª José Báguena: *Prevención y tratamiento de la malaria*; J. Vilardell: *Residuos nucleares*, *Hace...* y *Apuntes*; Luis Bou: *Pobreza y enfermedad*, *Acceso inalámbrico a la Red*, *La Tierra, una peonza con el corazón líquido*; Joandomènec Ros: *Evolución de los tetrápodos*; M.ª Rosa Zapatero: *Historia de la formación de las estrellas*; Angel González de Pablo: *La guerra del desierto*; Ramón Muñoz Tapia: *Taller y laboratorio*



Portada: Raúl Martín

SUSCRIPCIONES

Prensa Científica S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 934 143 344
Fax 934 145 413

Precios de suscripción:

	Un año	Dos años
España	65,00 euro	120,00 euro
Resto del mundo	90,00 euro	170,00 euro

Ejemplares sueltos:

El precio de los ejemplares atrasados es el mismo que el de los actuales.



ISSN 0210136X Dep. legal: B. 38.999 – 76

Imprime Rotocayfo-Quebecor, S.A. Ctra. de Caldes, km 3 - 08130 Santa Perpètua de Mogoda (Barcelona)

Printed in Spain - Impreso en España



HACE

...cincuenta años

LA EMANCIPACIÓN DE LA CIENCIA. «En la reunión anual de la Asociación Americana para el Progreso de la Ciencia, celebrada en Atlanta, el problema racial se planteó con toda crudeza. Cuatro de los afiliados a la AAPC se habían negado a participar. También algunos científicos negros destacados boicotearon el encuentro. A los negros que asistieron no se les privó de los actos sociales de la convención, pero quedaron excluidos de hoteles y restaurantes para blancos. Según constaba en parte de una resolución: 'Es necesario y deseable que todos los miembros puedan reunirse libremente para hablar de temas científicos, intercambiar ideas y difundir el estado actual del conocimiento... Tales objetivos no pueden cubrirse si la libre asociación de los miembros es impedida por obstáculos no naturales.'»

LOS NUEVOS JUBILADOS. «La idea de que la jubilación descontenta a las personas la contradice un estudio interdisciplinario que en la actualidad se lleva a cabo en la Universidad de Cornell. Tres cuartos de 700 varones recién retirados se declararon satisfechos con su nueva forma de vida; en general, gozaban de buena salud. En términos globales, los que se mostraron contentos con el retiro contaban con seguridad económica y habían hecho planes para el futuro.»

...cien años

CRIATURAS DEL CRETÁCEO. «En la ilustración se representan dos reptiles de gran tamaño disputándose la supremacía. Se trata de driptosaurios que, si bien harto menores que los tiranosaurios y de un período anterior, probablemente se asemejaban mucho a ellos en hábitos y aspecto general. La recreación se considera correcta, aunque no se cree que tales animales desplegaran tanta agilidad, pues ninguna de las huellas de pies de driptosaurio halladas lo sugieren. Las ilustraciones se basan en las réplicas de Charles R. Knight, que tan magistralmente complementan la exposición de muestras paleontológicas del Museo Americano de Historia Natural de Nueva York.»



Dinosaurios en acción, según Charles Knight, 1906.

UN UNIVERSO EN EXPANSIÓN. «Las fotografías de la bóveda celeste entera revelan del orden de cien millones de soles. Son cosa de nada. Los matemáticos saben que existe materia suficiente para formar treinta y dos mil millones de soles de la misma masa que el nuestro. La prueba la brinda la velocidad de los soles rápidos. Una diminuta fracción de esta materia aparece en forma de soles visibles. Pero, ¿y el resto? ¿Se halla en soles muertos, planetas y lunas? ¿Es posible que miles de millones de soles apagados vaguen por los baldíos del espacio? ¿Y todos rodeados de planetas muertos que aún giran midiendo el paso de unos años inanimados e inútiles?»

EL DÍA DE LOS ENAMORADOS. «Poca gente se da cuenta de la gran industria desarrollada al calor de las felicitaciones anuales del 14 de febrero. La idea parece proceder de Inglaterra; actualmente la practican los anglohablantes de todo el mundo. Alemania, aunque no reconoce la fiesta, suministra gran parte de las tarjetas y los artículos de regalo empleados en Estados Unidos. Pero en los últimos años, nosotros no sólo abastecemos a nuestro propio mercado, sino que exportamos grandes cantidades de tarjetas de felicitación a todo el mundo.»

...ciento cincuenta años

LOS PRECIOS DEL COMBUSTIBLE. «El combustible necesario para cocinar una cena en París cuesta casi tanto como la cena misma. El combustible escasea mucho y a los estadounidenses les sorprende hallar tiendas por toda la ciudad, provistas de estantes como de zapatería en los que se almacena madera cortada a trozos del tamaño de un dedo humano y agavillados como espárragos. Trozos de mayor tamaño se atan del mismo modo y se venden a unos precios astronómicos. La antracita, casi tan cara como la madera, se vende en cantidades minúsculas en esas tiendas de combustible.»

EL SINO DE LAS MOMIAS. "Las momias egipcias son a veces extraídas por los árabes para usarlas como combustible. Se trate de un faraón, sus esposas, sus sacerdotes o sus esclavos, se abren y trocean con la misma indiferencia que si fueran troncos de pino. Los bálsamos y pegamentos que se usaron para embalsamarlas las han convertido en un buen sustituto del carbón bituminoso; así, los mismos medios que sirvieron para preservarlas se han convertido en los agentes activos de su destrucción."

APUNTES

SALUD PUBLICA

OMS, DDT y mosquiteras

La Organización Mundial de la Salud (OMS) contribuyó a que el Convenio de Estocolmo sobre Contaminantes Orgánicos Persistentes, adoptado en 2001 y en vigor desde 2004, permitiese el uso sanitario del DDT. Quienes pese a ello siguen haciendo campaña en favor del DDT parece que se proponen que la OMS, que no le ha estado negando algún papel en la lucha contra el paludismo, le devuelva la primacía. A finales de 2005 se publicaba el Plan Estratégico Global 2005-2015 de la campaña *Roll Back Malaria*, promovida, entre otras instituciones y empresas, por la OMS y el Banco Mundial. Reiterando criterios ya conocidos, el Plan afirma que las mosquiteras tratadas con piretroides —insecticidas sintéticos biodegradables—, en especial las que no necesitan ser impregnadas de nuevo en 4 o 5 años, constituyen "la intervención prioritaria" a corto y medio plazo "en casi todas las áreas donde la malaria es endémica". La rociada de paredes con insecticida —el DDT entre otros— se recomienda para aquellas zonas —como ciertos países sudafricanos y del Cuerno de Africa— donde el paludismo es inestable y propenso a brotes epidémicos. Pero para el resto del Africa Subsahariana, donde es estable, la transmisión elevada y faltan medios para rociar a gran escala, las mosquiteras tienen "ventajas importantes": exigencias operativas algo menores y mayor posibilidad de concentración en los grupos de riesgo. (A veces se abogaba por el DDT por ser más barato, pero parece que en algunos casos la diferencia de precios por metro cuadrado de pared se ha ido esfumando; y según el cuestionario de la OMS sobre el DDT, si bien "se ha demostrado que en ciertos emplazamientos la aplicación del DDT es más económica que el uso de las mosquiteras tratadas", los costes operativos de emplearlo en poblaciones muy dispersas "pueden ser prohibitivos". En documentos previos al Plan se sostiene que la relación entre coste y eficacia de rociadas y mosquiteras es "similar" o que no cabe hacer "afirmaciones generalizadas".)

PLANETAS

El tigre de Encélado

Encélado, uno de los satélites de Saturno, está resultando uno de los objetivos más interesantes de la misión Cassini. Las observaciones del año pasado parecían indicar que penachos de vapor de agua y polvo se alzaban centenares de kilómetros por encima del hemisferio sur de la luna, alimentaban una tenue atmósfera y producían uno de los anillos de Saturno. La cámara del Cassini ya ha sorprendido a los penachos en acción. Parecen salir de unas "rayas de tigre", unas fisuras paralelas que fulguran en las imágenes infrarrojas, señal de que escapa calor. De ese modo Encélado se convierte en el tercer cuerpo del sistema solar (tras la Tierra e Io, luna de Júpiter) al que se le conoce actividad volcánica. Nada se sabe de la causa de su actividad geológica y de que, al contrario que el simétrico tigre de Blake, impresione por su "temible" asimetría entre los hemisferios norte y sur. El satélite es tan pequeño —sólo tiene 500 kilómetros de diámetro— que el calor se habrá fugado hace mucho tiempo de sus entrañas más profundas. No parece tampoco que las fuerzas de marea aporten tanta energía. Para más detalles habrá que esperar al próximo acercamiento de Cassini, en marzo de 2008.

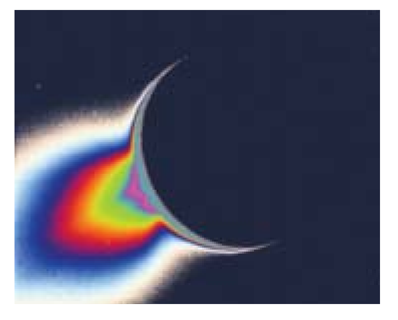


Imagen del penacho emergente de Encélado, coloreado para acentuar los contornos.

—George Musser

ENTOMOLOGIA

Las abejas reconocen rostros

Las abejas pueden aprender a reconocer rostros humanos. Adrian G. Dyer, de la Universidad de Cambridge, y sus colaboradores las han enseñado a asociar fotos de rostros en blanco y negro a un premio goloso (una bebida azucarada) o a una sanción agria (una solución de quinina). Cuando ya no mediaban ni premio ni castigo, revoloteaban a seis o siete centímetros de las fotografías antes de aterrizar del 80 al 90 por ciento de las veces junto a la "cara con premio". Tampoco lo hacían mal cuando entre las figuras a elegir había caras nuevas y retratos de palotes. Estos resultados demuestran que el reconocimiento de rostros, que podría parecer una aptitud neuronal compleja, no requiere de una gran potencia cerebral; las abejas tienen menos del 0,01 por ciento de las neuronas con que contamos nosotros.

—Philip Yam



Es él: una abeja adiestrada para reconocer caras revolotea ante una foto.

NASA/JPL/INSTITUTO DE CIENCIAS ESPACIALES (*arriba*); CORTESIA DE ADRIAN G. DYER, REIMPRESO CON PERMISO DE *JOURNAL OF EXPERIMENTAL BIOLOGY*, VOL. 208, 15 DE DICIEMBRE DE 2005 (*abajo*)

DIETETICA

Incertidumbres de la fibra

¿Cuál es la verdadera relación entre la fibra alimenticia y el cáncer colorrectal? Según el diseño del estudio y el tipo de fibra investigada, lo mismo se descubre que la fibra nos protege del cáncer colorrectal que se llega a la conclusión de que no tiene ningún efecto o incluso de que aumenta el riesgo. El último estudio a gran escala, publicado en el *Journal of the American Medical Association* del 14 de diciembre de 2005, reunía los resultados de 13 estudios prospectivos, que en conjunto siguieron a unas 725.000 personas durante 20 años. Aunque no pudieron descontar factores alimenticios superpuestos ni tomar en cuenta la solubilidad de la fibra, los autores averiguaron que los que menos fibra comían (los que no llegaban a 10 gramos diarios) corrían más peligro de contraer cáncer colorrectal. Pero las grandes ingestiones de fibra no reducían el riesgo. Los resultados dan también a entender que la fibra de los cereales y de los granos integrales, pero no la de las frutas y verduras, confieren una probabilidad algo menor de padecer cáncer colorrectal.

—*J. R. Minkel*



Las grandes ingestiones de fibra quizá no rebajen el riesgo de contraer cáncer colorrectal.

PEDAGOGIA

La guerra de los métodos de aprender a leer

Podría parecer que un debate científico acerca de cuál es el mejor modo de enseñar a leer resulta, pese a sus implicaciones sociales, demasiado abstruso para tener dimensiones políticas. El método fónico enseña el sonido de las letras y sílabas; a partir de esas unidades se aprende a leer palabras completas con independencia de su significado. Los métodos globales, en cambio, enseñan a leer palabras completas, asociadas al sentido. Los gobiernos británico y francés se han inclinado por el método fónico puro. El ministro francés de educación, Gilles de Robien, atribuye el fracaso escolar al método seguido hasta ahora en las escuelas. Propone el abandono de "todo método que guarde alguna relación, cercana o lejana, con los métodos globales o semiglobales". Y la ministra británica, Ruth Kelly, hace suyas las conclusiones de un informe que recomienda que sólo se use el método fónico. Las asociaciones británicas de maestros han mostrado su desacuerdo. Consideran que se necesitan varias estrategias, no una sola, y que en la lectura "hay algo más que la fonética". En Francia, André Ouzoulias recuerda el elevado porcentaje de alumnos que repetía el primer curso de primaria a mediados de los años sesenta, cuando se empleaba el método fónico; hubo que habilitar numerosas clases especiales para niños con dificultades de lectura.

METEOROLOGIA

¿A qué se debe la humedad?



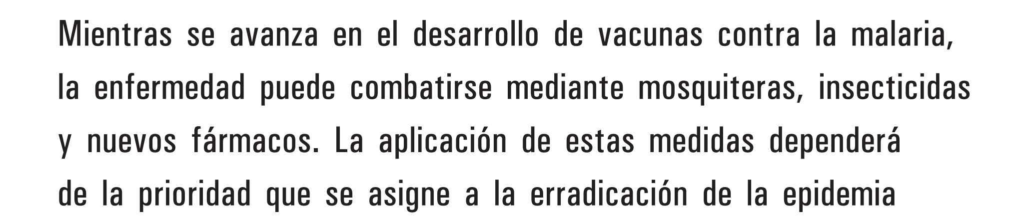
STOCKFOOD (*arriba*)

La humedad expresa la cantidad de agua existente en el aire, que varía de acuerdo con el interminable ciclo de la evaporación y la condensación. El agua de la superficie terrestre, de mares, ríos y lagos, pasa a la atmósfera en forma de vapor. El aire continúa absorbiendo agua hasta que ya no puede contener más. A partir de ese momento el vapor de agua se puede condensar; forma nubes y acaba regresando a la superficie en forma de lluvia o nieve. La humedad relativa da el tanto por ciento de vapor de agua que contiene la atmósfera con respecto al máximo que podría contener. El aire caliente puede contener, por unidad de volumen, más vapor de agua que el aire frío, pero a cualquier temperatura al aire que carga la mitad de su humedad máxima correspondiente se le asigna una humedad relativa del 50 por ciento. Los partes meteorológicos suelen mencionar este número porque nos sentimos incómodos cuando los valores elevados de humedad relativa van acompañados de temperaturas altas. La razón es que dependemos de la evaporación para refrescarnos. Cuando el aire ya se encuentra saturado de humedad, la piel no logra librarse del sudor. El punto de rocío del aire es la temperatura a la que se produce la saturación. Como ejemplo claro de este efecto vale lo que ocurre con una lata bien fría de refresco en un día de bochorno. Recordemos que el aire caliente puede albergar más agua que el frío. Cuando el aire que hay alrededor de la lata se enfría, alcanza su punto de rocío y deja agua condensada en el exterior del recipiente. En la atmósfera, el aire que alcanza su punto de rocío forma nubes; el agua del aire se fija sobre partículas de polvo. Cuando la temperatura de una cantidad suficiente de aire cae por debajo del punto de rocío, se crea una nube de agua y polvo. Pesa demasiado y se produce una precipitación.

—*Jeffrey Hovis*
Servicio Meteorológico Nacional (en Charleston, EE.UU.)



Prevención y tratamiento de la malaria



Mientras se avanza en el desarrollo de vacunas contra la malaria, la enfermedad puede combatirse mediante mosquiteras, insecticidas y nuevos fármacos. La aplicación de estas medidas dependerá de la prioridad que se asigne a la erradicación de la epidemia

Claire Panosian Dunavan

Hace tiempo en Gambia, en el Africa occidental, Ebrahim, un niño de dos años, estuvo a punto de morir de malaria. Decenios después, el doctor Ebrahim Samba recuerda lo ocurrido cuando se mira al espejo. Su madre, que había enterrado ya a varios hijos cuando él enfermó, le hizo un corte en la cara en un último y desesperado intento de salvarle la vida. El niño no sólo sobrevivió, sino que, además, llegó a convertirse en director regional de la Organización Mundial de la Salud para Africa.

¿Qué salvó a Ebrahim Samba? Por supuesto, nada que guardara relación con la escarificación ¿Fue acaso la cepa del parásito que le invadió el torrente circulatorio? ¿Su constitución genética o inmunitaria? ¿Su estado nutricional? Tras siglos de lucha contra la malaria y después de vencerla en una parte extensa del mundo, quedan todavía muchos interrogantes por despejar. ¿Qué determina que viva o muera un niño caído en las garras de esta plaga? Todavía hay esperanza. A partir del examen de supervivientes de la malaria y del seguimiento de otras pistas, los expertos trabajan en el desarrollo de vacunas. Otras medidas ocupan ya las primeras líneas de defensa: mosquiteras empapadas de insecticida y otros sistemas antimosquito, así como nuevos fármacos basados en una planta medicinal de la farmacopea tradicional china.

En los próximos años, el mundo necesitará todas las armas disponibles contra la malaria. Amén de causar muertes, esta enfermedad frena el desarrollo humano y económico. Combatirla constituye una prioridad internacional.

Un criminal en Africa

Cuatro especies principales del género *Plasmodium*, el parásito responsable de la malaria, infectan a los humanos. Al menos una de ellas infesta todos los continentes (excepto la Antártida), en mayor o menor grado. En la actualidad, sin embargo, el Africa subsahariana constituye no sólo la mayor reserva de *P. falciparum*, la especie más letal para los humanos, sino también el hogar de *Anopheles gambiae*, la más agresiva de la sesentena larga de especies de mosquito que transmiten la malaria a las personas. Cada año se producen en Africa quinientos millones de infecciones por *P. falciparum*, que se cobran entre uno y dos millones de muertos, sobre todo niños. En las zonas más castigadas, la malaria y sus complicaciones suponen entre el 30 y el 50 por ciento de los pacientes ingresados y más del 50 por ciento de los pacientes externos.

El cuadro clínico de la malaria por *P. falciparum*, en niños y adultos, no resulta tranquilizador. En una situación de máxima agresividad, a la fiebre y los escalofríos característicos les siguen una anemia grave, convulsiones y coma, fallo cardíaco y pulmonar, y la muerte. Quienes sobreviven padecen secuelas físicas, mentales y debilidad crónica. Aunque también se dan casos como el de Ebrahim Samba, quien, tras sufrir la enfermedad, no tuvo efectos secundarios. (Todavía hoy, más de medio siglo después de hallarse al borde de la muerte, el caso de Samba desconcierta a los expertos.)



1. CON LA PICADURA DE UN MOSQUITO INFECTADO comienza el ciclo letal de la malaria, una enfermedad que mata entre uno y dos millones de personas cada año, sobre todo niños del Africa subsahariana.

La anemia falciforme y otras alteraciones hereditarias de la hemoglobina frenan la infección a través del torrente sanguíneo. Se supone que los anticuerpos y las células inmunitarias que se desarrollan con el tiempo protegen a un gran número de africanos del ataque palúdico. Ebrahim Samba ofrece un buen ejemplo de esta transformación derivada de infecciones repetidas. Tras su temprano enfrentamiento con la muerte, no sufrió más crisis de malaria; hasta el día de hoy, no utiliza medidas preventivas para evitar nuevos ataques.

La historia de Samba abona la esperanza de que algún día quizá dispongamos de vacunas que produzcan las mismas defensas que aparecieron de forma natural en él, reduciendo las complicaciones y las muertes que causa la malaria en las regiones endémicas. Otro tipo de vacunas bloquearían la infección (al menos durante un breve período de tiempo) en turistas, cooperantes, militares en misiones de paz y otros visitantes que necesitaran una protección de menor duración.

Sin embargo, tampoco podemos sobrestimar el potencial de las vacunas. Los parásitos transmisores de la malaria revisten mayor complejidad que los virus y bacterias causantes de enfermedades para las que ya existen vacunas. Las vacunas antipalúdicas quizá no alcancen nunca la eficacia de las del sarampión o la polio, que protegen a más del noventa por ciento de quienes completan el tratamiento. Sin vacuna, la tragedia palúdica en Africa crecerá como una hidra de mil cabezas.

Al frente de la lista de los problemas actuales hallamos a las cepas de *P. falciparum* resistentes a las medicinas. Primero se extendieron por Sudamérica y Asia y luego por todo el continente africano. Le siguen la resistencia de los mosquitos a los insecticidas, la ausencia de infraestructuras en salud pública y la extrema pobreza, la principal traba para la prevención de las infecciones. Por último, la pandemia de sida que azota Africa compite por los recursos sanitarios y desaconseja las transfusiones de sangre en casos de anemia palúdica grave.

Pero no podemos perder la esperanza. Sería un error pensar que Africa nunca se liberará de las cadenas de la malaria. La historia económica, por ejemplo, nos enseña que no tiene por qué ser así.

Malaria y pobreza

La malaria no siempre ha estado confinada en los trópicos. En el siglo XX afectaba, entre otros, a Escandinavia y el Medio Oeste de los Estados Unidos. Los hechos que llevaron a la expansión de la epidemia hacia regiones templadas y, más recientemente, hacia extensas zonas de Asia y Sudamérica, guardan mayor relación con la pobreza que con la biología de la enfermedad.

Tomemos, por ejemplo, el recorrido de la malaria desde su último reducto en los EE.UU., el sur pobre y rural. El momento decisivo llegó con la Gran Depresión, cuando el ejército, la Fundación Rockefeller y la Autoridad del Valle de Tennessee (AVT) comenzaron el drenaje y la pulverización con petróleo de miles de áreas de reproducción del mosquito; ello se conjugó con el reparto de quinina (antimalárico de origen vegetal descubierto en Sudamérica) para purgar a las personas de los parásitos y con ello evitar su transmisión. Pero las medidas no se detuvieron aquí. Los ingenieros de la AVT que llevaron la energía hidroeléctrica al sur, regularon también el nivel de los pantanos para aislar las larvas de los mosquitos e instalaron mosquiteras en puertas y ventanas. A medida que la malaria retrocedía, crecía la economía local.

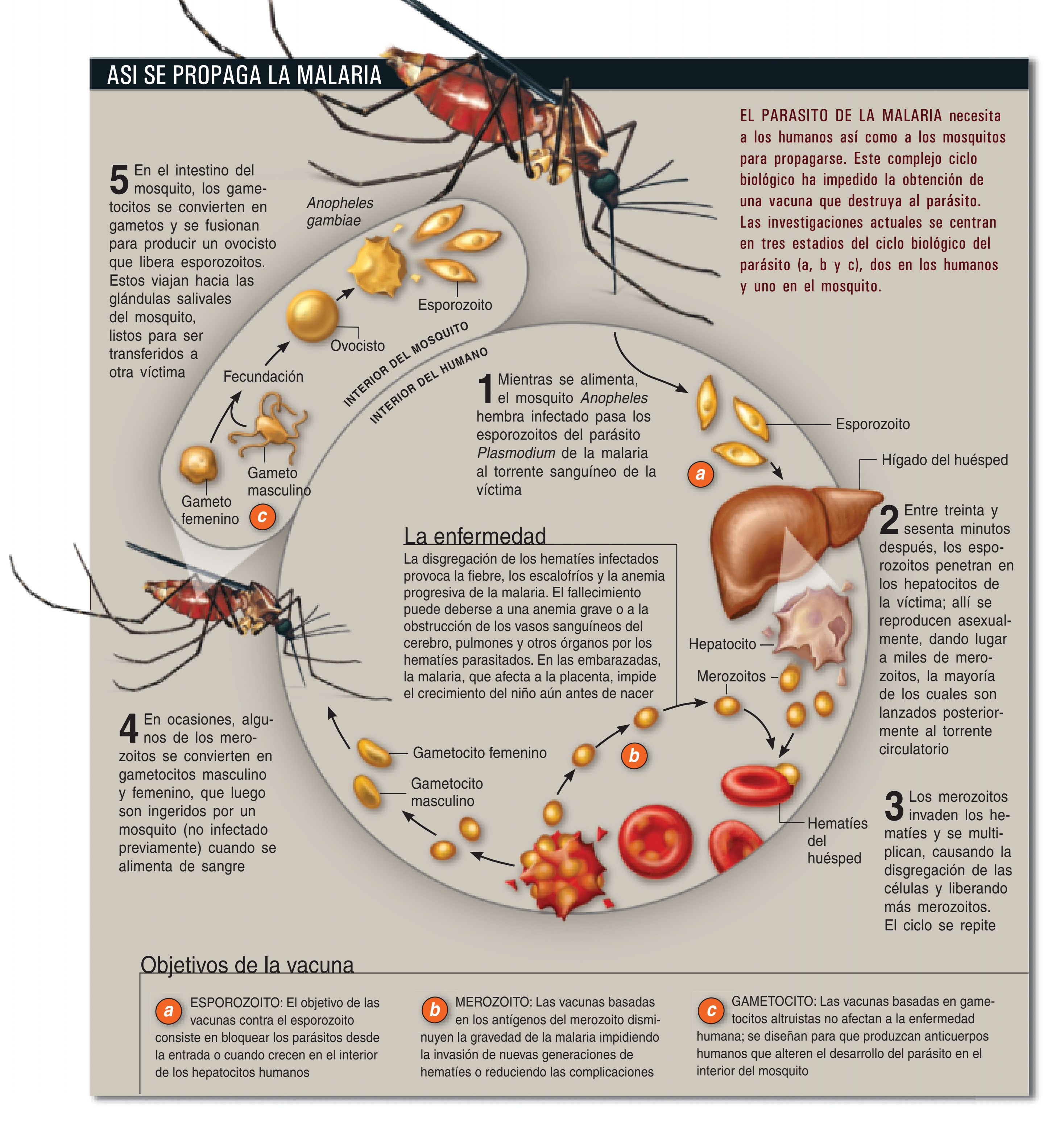
Llegaron luego los días dorados del DDT (diclorodifeniltricloroetano). En plena Segunda Guerra Mundial, la fuerza aérea fumigó con el insecticida las zonas de malaria del Pacífico. Las autoridades sanitarias tomaron el relevo. Cinco años después, la fumigación selectiva del interior de las casas se convertía en la piedra angular de la erradicación mundial de la malaria. En 1970, la fumigación con DDT, la eliminación de las áreas de reproducción del mosquito y el uso generalizado de fármacos antipalúdicos salvó a más de quinientos millones de personas, casi un tercio de los que habían vivido bajo la amenaza de la epidemia.

Africa subsahariana constituía un caso especial. Si exceptuamos algunos esfuerzos provisionales, allí nunca se llevaron a cabo programas sistemáticos de erradicación. No obstante, la cloroquina (compuesto semejante a la quinina, sintético y barato, introducido tras la segunda Guerra Mundial) permitió a países con escasos recursos sustituir las operaciones de fumigación, complejas y extensas, por la acción de una red sanitaria. Mediante la distribución de pastillas a los afectados por fiebre se salvaron millones de vidas en los años sesenta y setenta del siglo pasado. Pero, poco a poco, la cloroquina comenzó a fallar frente al *P. falciparum* palúdico. Dada la insuficiencia de infraestructuras y conocimientos para hacer frente al ataque de los mosquitos vectores africanos, se produjo un rebrote de la mortalidad.

Malaria y pobreza parecen compañeras inseparables. Hoy en día, en numerosos hogares africanos la malaria no sólo limita los ingresos y detrae fondos destinados a la alimentación o la escolarización de los más jóvenes, sino que estimula la natalidad, pues las familias de las víctimas asumen que la enfermedad siempre les arrebatará hijos. A escala regional, la epidemia ahuyenta inversiones extranjeras, turismo y comercio. Al continente le cuesta más de doce mil millones de dólares anuales, el cuatro por ciento

Resumen/Combatir la malaria hoy

- Los expertos trabajan en el desarrollo de vacunas que prevengan la malaria o disminuyan su gravedad.
- En la lucha contra la enfermedad se recurre al uso de mosquiteras tratadas con insecticidas, la fumigación del interior de las casas y nuevas combinaciones farmacológicas basadas en una planta medicinal de la farmacopea tradicional china.
- Pero el sida y otros flagelos compiten por los fondos que podrían destinarse a la lucha contra la malaria. La aplicación de dichas medidas dependerá, por tanto, de la prioridad que se asigne a la erradicación de la epidemia.



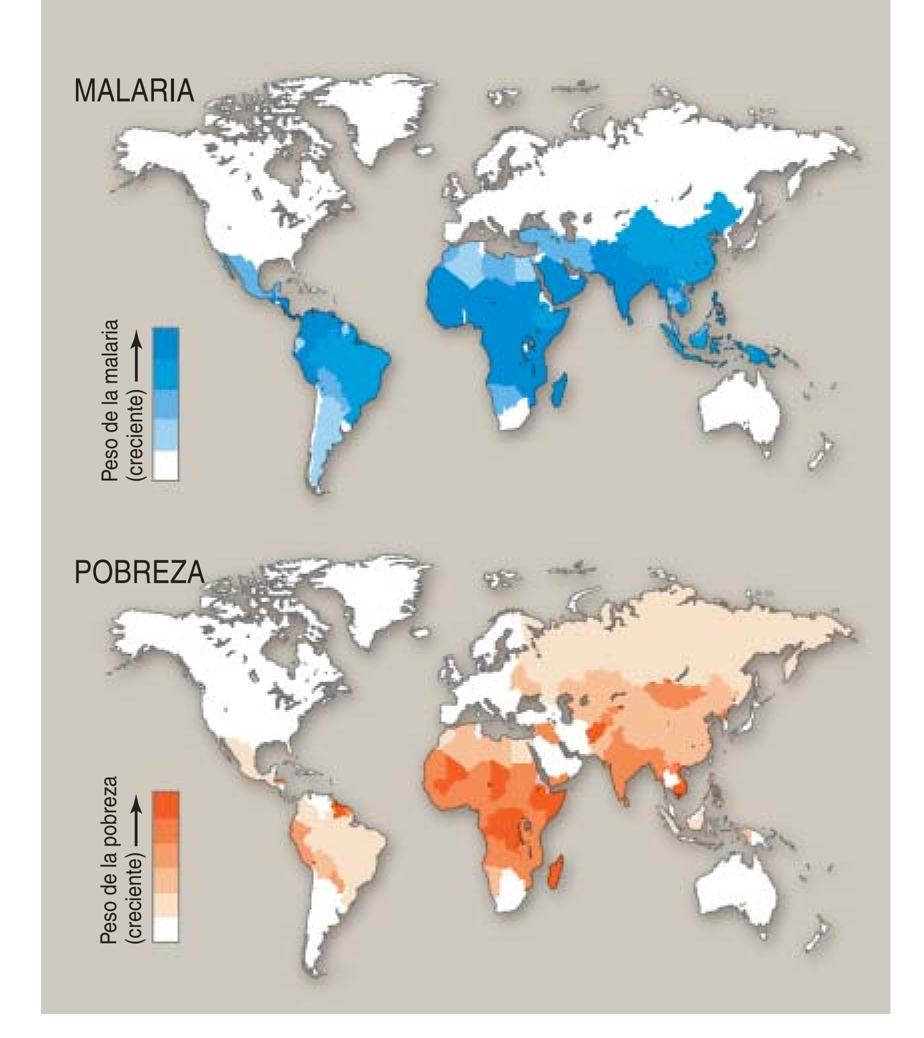
JEN CHRISTIANSEN

del producto interior bruto africano. En resumen, en muchos lugares el paludismo permanece atrincherado debido a la pobreza, al propio tiempo que la crea y perpetúa.

Medidas antimosquito

Existe todavía un pavoroso desconocimiento en torno al mecanismo de infección de la malaria. El hombre la contrae a través de las picaduras nocturnas de *Anopheles*, el mosquito que hospeda al parásito. Sin embargo, algunos habitantes de zonas infestadas creen que la enfermedad la provoca un espíritu maligno o determinados alimentos. Ello pone de manifiesto otra necesidad apremiante: la educación de la población.

Mucho antes de que Ronald Ross y Giovanni Battista Grassi descubrieran, a finales del siglo XIX, que la malaria la transmitían los mosquitos, se habían empezado a idear sistemas para evitar las picaduras. Cinco siglos antes de la era cristiana, Herodoto describió en su obra *Historias* que los egipcios de tierras pantanosas se protegían con las redes de pesca: “Cada hombre tenía una red que utilizaba durante el día para pescar, pero por la noche le daba otro uso: la extendía sobre la cama... Los mosquitos pican



2. MALARIA Y POBREZA abarcan territorios comunes. A los costes de prevención y tratamiento hemos de sumar la pérdida de ingresos, de inversiones extranjeras y de riqueza generada por el turismo. El crecimiento económico anual de los países con malaria endémica rondó el 0,4 por ciento del Producto Interior Bruto per cápita entre 1965 y 1990, comparado con el 2,3 por ciento del resto del mundo.

JEN CHRISTIANSEN; FUENTE: ROLL BACK MALARIA

a través de una manta o de una colcha de lino... pero, a través de la red, ni lo intentan". Basados en este pasaje, algunos defensores de las mosquiteras consideran que estas redes empapadas de aceite de pescado constituyeron las primeras ropas impregnadas de repelente de la historia.

La combinación de productos textiles e insecticidas no llegó, a buen seguro, hasta la Segunda Guerra Mundial. El ejército estadounidense destinado al Pacífico Sur comenzó a empapar mosquiteras y hamacas con DDT al cinco por ciento. Luego, a consecuencia de las múltiples manifestaciones públicas en contra del DDT, las mosquiteras pasaron a tratarse con insecticidas biodegradables: los piretroides. Lo que significó un gran avance.

En 1991, la primera aplicación a gran escala de mosquiteras tratadas con piretroides junto con fármacos antipalúdicos redujo a la mitad la mortalidad de los niños menores de cinco años en Gambia. En ensayos posteriores, sin fármacos, se obtuvieron resultados similares en cuanto a supervivencia en Ghana, Kenia y Burkina Faso, junto a una mejora sustancial en la salud de las embarazadas. Es más, a medida que se extendía su uso, familias y comunidades enteras se beneficiaron de las mosquiteras (incluso las personas que no dormían bajo ellas).

Ese sistema dual de protección tenía sus puntos débiles. Las mosquiteras impregnadas de insecticida resultaban eficaces sólo si los mosquitos transmisores de la malaria picaban en el interior de las viviendas durante las horas de sueño, una conducta en modo alguno universal. Además, las mosquiteras, incómodas, dan calor a los que duermen bajo ellas. Hasta hace poco, mientras no se disponía de PermaNet y Olyset (dos mosquiteras tratadas con piretroides de larga duración), había que reimpregnar las mosquiteras cada seis o doce meses para garantizar su eficacia. Si sumamos su coste, de entre 1,7 y 5 euros la unidad, las mosquiteras con o sin insecticida quedan simplemente lejos del alcance de la población. Un estudio reciente llevado a cabo en Kenia demostró que sólo el 21 por ciento de las viviendas cuentan con, al menos, una mosquitera, de las cuales el seis por ciento están tratadas con insecticida. De un resumen de treinta y cuatro estudios realizados entre 1999 y 2004 se desprendía una conclusión más desoladora: sólo el tres por ciento de los africanos más jóvenes estaba protegido por mosquiteras impregnadas de insecticida. Pese a todo, su utilización parece ir en aumento.

También la resistencia a los insecticidas minaría la eficacia a largo plazo de las mosquiteras. En varias localidades keniatas y sudafricanas han aparecido variedades de mosquitos que inactivan los piretroides; algunos anofelinos están retardando su muerte por contacto con piretroides, una conducta adaptativa preocupante denominada desprendimiento inmediato (o kdr, de "knockdown resistance").

Dada la escasez de nuevos insecticidas de uso médico (sobre todo porque apenas se invierte en su desarrollo), una solución consistiría en tratar las mosquiteras con otros insecticidas de uso agrícola. Descifrar las señales olfatorias que atraen a los mosquitos hacia los humanos podría abrir nuevas vías en la obtención de repelentes. En cruel contrapartida, se acaba de publicar que un cambio del olor corporal debido a la presencia de *P. falciparum* en la sangre podría atraer a los mosquitos: los escolares de Kenia que albergan gametocitos (la fase del parásito palúdico que pasa al mosquito cuando éste pica a la víctima humana) sufren el doble de picaduras que sus compañeros no infectados.

¿Y si se utilizaran estas criaturas aladas para matar a los parásitos de la malaria? En teoría, podríase aplicar la ingeniería genética para impedir la multiplicación del parásito antes de que los protozoos abandonaran las glándulas salivales del insecto. Si los mosquitos transgénicos lograran desplazar a sus parientes no mani-

pulados en la naturaleza, evitarían la difusión de los parásitos entre las personas. En fecha reciente, se han identificado genes nativos que impiden la multiplicación de la malaria en los mosquitos *Anopheles*. Y se están esbozando proyectos de manipulación genética de cepas de diversas especies. Sin embargo, una vez obtenidos en el laboratorio estos "insectos de Troya", su introducción en el mundo real planteará nuevas cuestiones, incluidas las éticas.

Por ahora, la tradicional fumigación del interior de las viviendas con DDT sigue constituyendo una medida fundamental para la salud pública en muchos lugares, dentro y fuera de Africa. Aplicado a superficies, el DDT permanece activo durante seis meses o más. Reduce el contacto entre los humanos y el mosquito mediante dos mecanismos clave: repele a unos mosquitos antes de que entren en la vivienda y mata a otros que, después de comer, se posan sobre las paredes impregnadas. En 1999 y 2000, su eficacia quedó ampliamente demostrada en KwaZulu-Natal. Un *A. funestus* resistente a los piretroides y a otros fármacos había causado el mayor número de casos de *P. falciparum* desde que la provincia sudafricana puso en marcha su programa de control de la malaria. La reintroducción de la fumigación con DDT junto con nuevos fármacos logró en dos años una disminución del 91 por ciento de la prevalencia.

Tratamientos médicos

Las medidas antimosquitos no pueden, por sí solas, ganar la guerra contra la malaria. Se requieren fármacos más eficaces y atención médica para los millones de jóvenes y adultos que cada año caminan por la cuerda floja palúdica. Algunos confían en los herbolarios del poblado y los curanderos ambulantes. Otros toman píldoras de procedencia, calidad y eficacia desconocidas, o falsificadas, que familiares o vecinos obtienen de fuentes irregulares. En Africa, el se-

DDT: un insecticida convertido en símbolo

Una campaña mundial de erradicación de la malaria basada en la fumigación de las casas con DDT (diclorodifeniltricloroetano) logró en los años cincuenta grandes éxitos. En menos de dos décadas, el plaguicida permitió a muchos países controlar la enfermedad. En la India, por ejemplo, las muertes por malaria cayeron en picado: de ochocientas mil a casi cero.

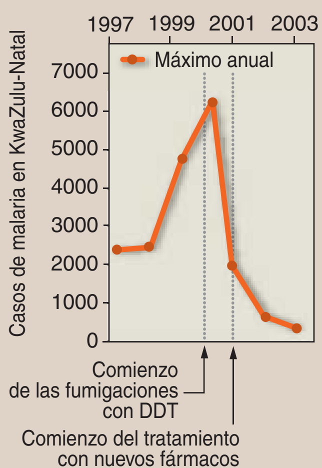
Luego, en 1972, el gobierno estadounidense prohibió utilizarlo en la fumigación de cosechas, aunque se permitió su uso médico y alguna otra aplicación de menor importancia. Se dijo que el elocuente libro de Rachel Carson, *Silent Spring* (Primavera tranquila) había precipitado la prohibición. De forma meticulosa, Carson rastreó el camino seguido por el DDT a través de la cadena alimentaria. Mostró que, a medida que se iba acumulando, exterminaba o alteraba genéticamente insectos y otros animales. El DDT se convirtió en un símbolo del peligro de jugar a ser Dios con la naturaleza. Los países desarrollados, que habían conseguido librarse de la malaria dentro de sus fronteras, abandonaron la utilización del plaguicida. La mayor parte de Europa imitó a los EE.UU. y prohibió su uso agrícola en los años setenta.

Para el Africa subsahariana, en donde la malaria aún arrasa, estas decisiones han supuesto la pérdida de un arma valiosa. La mayoría de los países no recurren al DDT, no porque lo haya prohibido su gobierno (de hecho, la utilización sanitaria del insecticida está permitida en la mayoría de las regiones del planeta en donde la malaria es endémica), sino porque los países ricos y las organizaciones donantes se resisten a financiar proyectos que incluyan fumigaciones con DDT, incluso de forma responsable.

Muchos de los expertos en malaria piensan que el DDT debería reconsiderarse. Amén de resultar tóxico para los mosquitos, señalan, los ahuyenta de las paredes y puertas fumigadas antes de que piquen e impide su entrada a la vivienda. Tóxico, irritante y repelente, todo en uno. Además, dura dos veces más que las otras alternativas y cuesta la cuarta parte que el siguiente insecticida más barato.

La trayectoria letal del compuesto a través de la cadena alimentaria hunde sus raíces en la fumigación agrícola masiva (sobre todo de los campos de algodón), no en su utilización, mucho más moderada, en las viviendas. El tratamiento de cien hectáreas de campos de algodón requiere unos mil cien kilogramos de DDT durante cuatro semanas. La fumigación del interior de una casa, en cambio, se completa con apenas medio kilogramo, aplicado una o dos veces al año.





LOS CASOS DE MALARIA DISMINUYERON drásticamente en KwaZulu-Natal cuando el gobierno sudafricano fumigó las viviendas con DDT y luego administró a los pacientes un tratamiento combinado basado en la artemisina (*gráfico*). Sudáfrica, uno de los pocos países africanos con recursos suficientes para costear su propio programa antipalúdico, no tiene que recurrir a la ayuda de países reacios a la utilización del DDT. Los aleros de las casas africanas típicas, como los que aparecen en la fotografía, facilitan la entrada de los mosquitos.

El DDT, por sí solo, no salvará al mundo de la malaria. La fumigación de las casas, por ejemplo, resulta efectiva sólo contra los mosquitos que pican en su interior. Se requieren fármacos para el tratamiento de pacientes infectados, así como nuevas medidas antimosquito. No obstante, la mayoría de los profesionales de la salud que luchan contra la malaria consideran el uso dirigido del DDT una herramienta importante de su equipo de trabajo.

JEN CHRISTIANSEN; FUENTE: INFORME IOM (*gráfico*); KAREN DUTHIE *AfricanPictures.net* (*fotografía*)

tenta por ciento de los antimaláricos se compran en un mercado marginal de vendedores ambulantes, no en los hospitales ni en las farmacias.

A pesar de que su eficacia desciende en picado, la cloroquina se mantiene a la cabeza de los fármacos antipalúdicos consumidos en Africa. Le sigue la sulfadoxina-pirimetamina, un antibiótico que obstaculiza la síntesis de ácido fólico por el parásito. Por desgracia, las cepas de *P. falciparum* en Africa y otros lugares están empezando a eludir este compuesto conforme experimentan mutaciones secuenciales que, andando el tiempo, minarán la eficacia del fármaco.

Ante el espectro amenazante de la resistencia a los medicamentos, ¿podemos aprovechar los conocimientos sobre otras enfermedades infecciosas para reforzar el tratamiento farmacológico de la malaria? En los últimos decenios, la aparición de cepas resistentes de los gérmenes responsables de la tuberculosis, la lepra y el sida ha obligado a introducir cambios en las pautas terapéuticas.

Para prevenir la aparición de "superpatógenos", los tratamientos actuales combinan dos o tres fármacos. Se da por cierto que los tratamientos combinados combaten también la resistencia del *P. falciparum* de la malaria, sobre todo si incluyen una variedad de la *Artemisia annua*, una planta medicinal utilizada como antitérmico en la farmacopea de la China clásica. Los fármacos derivados de la *Artemisia* (o artemisinas) repelen los parásitos de la malaria más rápidamente que cualquier otra sustancia; además, bloquean la transmisión de humanos a mosquitos.

Ante tales ventajas se especula si la combinación de éstos con otros medicamentos antipalúdicos no prevendría o retrasaría la resistencia a la artemisina, no sólo en Africa sino también en el resto del mundo. Después de todo, nada garantiza que la malaria no vuelva algún día a sus antiguas moradas. Constituye una amenaza para el viajero global. En los últimos años, mosquitos infectados de *P. falciparum* han viajado como polizones en vuelos internacionales, infectando a inocentes espectadores situados a pocos kilómetros de los aeropuertos, lejos del medio natural de la malaria.

Sin embargo, los nuevos remedios combinados presentan un problema: su coste, entre diez y veinte veces superior al de los fármacos antipalúdicos más populares —y cada vez menos eficaces— en Africa, resulta desalentador para la mayoría de las víctimas y los países afectados. Aun cuando los nuevos cócteles resultaran más baratos, las reservas mundiales de artemisinas no satisfarían la demanda. Además, deben recaudarse fondos para activar el ciclo de producción de las plantas (cultivo, cosecha y procesado), de unos 18 meses de duración. Novartis, el primer productor autorizado por la OMS para fabricar un tratamiento combinado de artemisina coformulado (artemeter más lumefrantina), puede que no disponga de la financiación y la materia prima necesarias para enviar ni siquiera una parte de los 120 millones de tratamientos que tenía previsto distribuir en 2006.

Pese a tanta dificultad, se espera conseguir, en un plazo de entre cinco y diez años, fármacos sintéticos y más baratos, con las propiedades químicas de las artemisinas vegetales (un enlace peróxido incrustado en un anillo). Cierto prototipo que se sintetizó a finales de los noventa comenzó a someterse a ensayos clínicos en 2004. Otra técnica prometedora, que sortearía la extracción botánica y la síntesis química, se basa en empalmar genes de *A. annua* y de levadura en *Escherichia coli* para luego obtener fármacos a partir de la bacteria modificada. Este método se desarrolló en la Universidad de California en Berkeley.

También la prevención, frente al tratamiento, de la malaria está ganando adeptos entre niños y mujeres embarazadas de Africa, los huéspedes más vulnerables. Cuarenta años atrás se demostró que dosis bajas de antimaláricos administradas como profilaxis a las embarazadas nigerianas aumentaban el peso de los recién nacidos. Ahora se ha sustituido ese tratamiento por una serie de tomas de sulfadoxin-pirimetamina durante el embarazo, la infancia y, en dosis mayores, aprovechando las vacunaciones infantiles.

La fórmula anterior reduce las infecciones y la anemia; pero una vez la resistencia se extienda por Africa, ¿qué tratamiento preventivo sustituirá a la sulfadoxin-pirimetamina? Resultaría lógico pensar en las artemisinas. Sin embargo, estos agentes no sirven para la prevención porque su concentración en sangre disminuye muy rápidamente. Además, dosis repetidas de artemisina en mujeres y niños asintomáticos (una práctica aún no analizada) podrían producir efectos secundarios desconocidos. La prevención ideal se lograría, sin duda, mediante una vacuna.

Vacunas

A pesar de los progresos logrados en el curso de los últimos decenios, la creación de vacunas eficaces contra la malaria ha encontrado múltiples obstáculos. El principal: el intrincado ciclo biológico del parásito. Abarca varios estadios en los mosquitos y en los humanos. Y así, una vacuna podría matar un estadio, pero permitir el crecimiento de otro.

Otra dificultad proviene de la compleja dotación genética del patógeno. De las 5300 proteínas codificadas por el genoma del *P. falciparum*, menos del diez por ciento activan respuestas protectoras en los individuos expuestos de forma natural; pero, ¿cuáles?

Por si fuera poco, una buena respuesta frente a la vacunación antipalúdica exige la coordinación de varios brazos del sistema inmunitario humano: anticuerpos, linfocitos y el bazo, como motor de arranque. Sin embargo, tales respuestas no siempre se desarrollan ni siquiera en personas sanas, por no hablar de poblaciones ya expuestas a la malaria y otras enfermedades.

Hasta la fecha, la mayoría de las vacunas experimentales frente al *P. falciparum* se han polarizado en uno de los tres estadios biológicos del parásito (esporozoito, merozoito y gametocito). Se están diseñando también vacunas multifásicas, que ofrecerían mayor eficacia. La posibilidad de atacar a los esporozoitos (la fase del patógeno que el mosquito suele inocular a sus víctimas por medio de su probóscide) se empezó a vislumbrar en los años setenta, cuando expertos de la Universidad de Maryland hallaron que esporozoitos de *P. falciparum* debilitados con rayos X protegían a voluntarios humanos, aunque por un breve lapso de tiempo. Posiblemente la vacuna inducía que el sistema inmunitario neutralizase la entrada de los parásitos antes de que emigraran, una hora



3. EBRAHIM SAMBA, antiguo director regional de la OMS para Africa, conserva aún unas tenues señales de incisiones en sus mejillas. Se las hicieron a los dos años, cuando estuvo a punto de morir de malaria.



después, hacia su estación siguiente, el hígado.

La demostración de que los anticuerpos obtenidos artificialmente contra los esporozoitos ayudarían a mantener a raya a la malaria alentó la prosecución de las investigaciones. Tres decenios después, en 2004, se recogían los primeros frutos: una vacuna contra los esporozoitos redujo a más de la mitad los episodios graves de malaria en dos mil niños de entre uno y cuatro años en las zonas rurales de Mozambique, edad en que la población infantil africana se halla más expuesta a esta enfermedad. La fórmula usada en este ensayo clínico, el más prometedor hasta la fecha, llevado a cabo por un equipo del Centro de Salud Internacional del Hospital Clínico y la Universidad de Barcelona, incluía múltiples copias de un fragmento de proteína del esporozoito del *P. falciparum* fusionado con proteína del virus de la hepatitis B, añadida para aumentar su potencia. Con todo, los sujetos necesitaban tres inmunizaciones y el período de protección era de sólo seis meses.

Una versión mejorada de la vacuna RTS,S (o cualquiera de sus casi tres docenas de hermanas que se encuentran en fase de desarrollo) podría tardar en llegar al mercado unos diez años y a un precio prohibitivo. Por eso se han organizado varios proyectos, como la iniciativa para la vacuna de la malaria de Seattle, con participación del sector público y el privado, dedicados a la recaudación de fondos para los ensayos clínicos en curso.

Hay que tener en cuenta otro factor a propósito de las vacunas contra la malaria. Aun cuando se dispusiera ya de ellas, seguirían siendo imprescindibles los tratamientos y las medidas antimosquitos. Las vacunas conferirán una protección total. Además, puede que parte de la población expuesta a la enfermedad, sobre todo en el Africa rural pobre, no tenga acceso a las inoculaciones. Por tanto, al menos en un futuro próximo, habrá que contar con todas las medidas preventivas y curativas.

Fondos contra la malaria

La malaria sigue conformando un escenario desolador: se cobra aún al menos un millón de víctimas cada año, al propio tiempo que frena el desarrollo humano y económico. Ha llegado la hora de contraatacar. A lo largo de los últimos diez años se han dado pasos significativos. En 1998, la OMS y el Banco Mundial establecieron el programa Roll Back Malaria. En 2000, el G8 declaró la malaria una de las tres pandemias que esperaba domeñar, si no erradicar. La ONU creó el Fondo Mundial de lucha contra el sida, la tuberculosis y la malaria; se comprometió a detener e invertir la tendencia alcista de la malaria en un plazo de quince años. En 2005, el Banco Mundial puso de nuevo a la malaria en su punto de mira; el presidente George W. Bush anunció una ayuda quinquenal de 1200 millones de dólares para luchar contra la malaria en Africa mediante mosquiteras tratadas con insecticidas, la fumigación del interior de las casas y el uso de tratamientos farmacológicos combinados.

En fecha más reciente, el Banco Mundial ha empezado a estudiar la subvención de los tratamientos combinados con artemisina. En octubre de 2005, la Fundación Bill y Melinda Gates anunció tres ayudas de un total de 258,3 millones de dólares para financiar proyectos de investigación sobre vacunas contra la malaria, nuevos fármacos y medidas antimosquitos más eficaces.

Pese a todo el montante aportado es insuficiente para acometer las tareas pendientes. Al tiempo que se anunciaba la donación de la Fundación Gates, se presentaban los resultados de un estudio que analizaba la inversión mundial en investigación y desarrollo sobre malaria: en 2004, sólo se invirtieron 323 millones de dólares, una cantidad muy inferior a los 3200 millones de dólares anuales proyectados para reducir las muertes por malaria a la mitad en 2010.

Quizá haya llegado el momento de movilizar no sólo a los expertos sino también a la población general. Diez euros, lo que puede costarnos el menú del almuerzo, suponen ya un largo trecho para conseguir una mosquitera tratada con insecticida o un ciclo de tres días de tratamiento combinado con artemisina para un niño africano.

La autora

Claire Panosian Dunavan, especialista en medicina tropical, da clases en la facultad de medicina David Geffen de la Universidad de California en Los Angeles. Acaba de codirigir el informe *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance*.

Bibliografía complementaria

WHAT THE WORLD NEEDS NOW IS DDT. Tina Rosenberg en *New York Times Magazine*, págs. 38-43; 11 de abril de 2004.

SAVING LIVES, BUYING TIME: ECONOMICS OF MALARIA DRUGS IN AN AGE OF RESISTANCE. COMMITTEE ON THE ECONOMICS OF ANTIMALARIAL DRUGS, BOARD ON GLOBAL HEALTH. Dirigido por Kenneth J. Arrow, Claire B. Panosian y Hellen Gelband. Institute of Medicine. National Academies Press, 2004.

EFFICACY OF THE RTS,S/AS02A VACCINE AGAINST PLASMODIUM FALCIPARUM INFECTION AND DISEASE IN YOUNG AFRICAN CHILDREN: RANDOMISED CONTROLLED TRIAL. P. L. Alonso, J. Sacarlal, J. Aponte, A. Leach, E. Macete, J. Milman, I. Mandomando, B. Spiessens, C. Guinovart, M. Espasa, Q. Bassat, P. Aide, O. Fori-Anyinam, M. Navía, S. Corachan, M. Ceuppens, M. C. Dubois, M. A. Demoitie, F. Dubovsky, C. Menéndez y N. Tornieporth en *Lancet*, vol. 364, págs. 1411-1420, octubre 2004.

RESIDUOS



Pese a la inquietud pública que despierta la energía nuclear, está ganando cada vez más adeptos la idea de que podría ser el menos dañino de los procedimientos que generan grandes cantidades de electricidad. Varios países —Brasil, China, Egipto, Finlandia, India, Japón, Pakistán, Rusia, Corea del Sur y Vietnam— están construyendo o proyectando centrales nucleares. Pero esa tendencia aún no se ha extendido a EE.UU., donde desde hace unos treinta años no se han puesto los cimientos de ninguna central nuclear; tampoco ha llegado a Europa occidental.

Desarrollada de manera adecuada, la energía nuclear sería, en verdad, un recurso sostenible, inagotable a todos los efectos. Su explotación no contribuiría al cambio climático. Una técnica nuclear hasta cierto punto nueva salvaría los principales inconvenientes de los métodos actuales: el miedo a los accidentes nucleares, la posibilidad de que el combustible se desvíe hacia la fabricación de armas, la gestión de los peligrosos residuos radiactivos de larga duración y la merma considerable de las reservas mundiales de uranio rentable. La nueva técnica combinaría dos innovaciones: el tratamiento pirometalúrgico (el reciclado a alta temperatura de los residuos de los reactores) y los reactores avanzados de neutrones rápidos, capaces de fisionar los núcleos atómicos del combustible generado por el reciclado. La radiactividad de los residuos generados descendería en pocos siglos a niveles no peligrosos; desaparecería así la necesidad de guardarlos durante decenas de milenios.

NUCLEARES

Los reactores de neutrones rápidos extraerían mucha más energía del combustible nuclear reciclado, reducirían el riesgo de proliferación de armas nucleares y acortarían el tiempo en que habrá que aislar los residuos

William H. Hannum, Gerald E. Marsh y George S. Stanford



JANA BRENNING

Para que los neutrones provoquen una fisión nuclear eficiente, deben moverse o despacio o muy deprisa. Aunque fisionable y fisible se consideren palabras sinónimas ("que se puede fisionar"), conviene ceñir el adjetivo "fisible" sólo a los combustibles nucleares que pueden fisionarse con neutrones lentos. La mayoría de las plantas nucleares existentes cuentan con reactores de los llamados térmicos: funcionan con neutrones de velocidad (o energía) baja, que rebotan en el interior de los núcleos; utilizan, pues, material fisible. Los reactores térmicos generan calor y, por lo tanto, electricidad con un rendimiento bastante bueno, pero no minimizan la producción de residuos radiactivos.

En todos los reactores, la energía se genera al desintegrarse átomos de metales pesados (de gran masa atómica), sobre todo de uranio o de elementos derivados del uranio. El uranio natural es una mezcla de dos isótopos, el uranio 235, fácilmente fisible, y el uranio 238, mucho más estable.

La fisión del uranio la inician y sostienen neutrones. Cuando sobre el núcleo de un átomo fisible incide un neutrón, en especial un neutrón lento, probablemente se fisionará, liberando grandes cantidades de energía y más neutrones. De esos neutrones emitidos, algunos chocarán entonces con otros átomos vecinos y estos se rompen. Se crea así una reacción nuclear en cadena. El calor correspondiente se extrae del reactor; fuera ya de éste, transforma agua en vapor, que mueve una turbina; ésta acciona un generador eléctrico.

El uranio 238 no es fisible pero sí fisionable: a veces se desintegra cuando lo golpea un neutrón rápido. De él también se dice que es "fértil", porque cuando un átomo de uranio 238 absorbe un neutrón y no se desintegra, se transmuta en plutonio 239; lo mismo que el uranio 235, el isótopo de plutonio es fisible y puede alimentar una reacción en cadena. Tras unos tres años de servicio, el plutonio contribuye con más de la mitad de la energía que genera la planta. Corresponde al tiempo que transcurre antes de que se extraiga el combustible de un reactor actual, operación necesaria a causa del agotamiento del uranio 235 y de la degradación debida a la radiación.

En un reactor térmico, los neutrones, que nacen rápidos, se refrenan (o "moderan") al interaccionar con los

átomos ligeros vecinos, tales como el hidrógeno del agua que fluye por los núcleos de los reactores. Los 440 reactores nucleares comerciales en servicio son todos térmicos, salvo dos, y en su mayoría, incluidos los 103 reactores productores de corriente estadounidenses y los nueve españoles, emplean agua tanto para moderar los neutrones como para transportar el calor a los generadores de electricidad asociados: se los llama reactores de agua ordinaria.

En las centrales nucleares, cuando "se quema" el uranio, se consumen átomos del metal pesado. Aun cuando la planta empiece con un combustible cuyo contenido en uranio 235 se haya enriquecido, al cabo de unos tres años la mayor parte de ese uranio fisionado habrá desaparecido. Cuando los técnicos extraen el combustible empobrecido sólo se ha gastado aproximadamente la vigésima parte de los átomos potencialmente fisionables que contiene (uranio 235, plutonio y uranio 238); el combustible agotado aún encierra del orden del 95 por ciento de su energía original. Además, sólo aproximadamente el diez por ciento del mineral de uranio se convierte en combustible mediante el proceso de enriquecimiento. Durante el mismo, aumenta considerablemente la concentración de uranio 235; en las centrales hoy en servicio se aprovecha, pues, menos de una centésima parte del contenido energético total del mineral.

El combustible usado procedente de los reactores térmicos actuales aún posee, por lo tanto, la capacidad de alimentar mucha combustión nuclear. Como las reservas mundiales de uranio son finitas y un ininterrumpido aumento del número de reactores térmicos acabaría con el uranio barato disponible, sería irracional desechar ese combustible agotado o los "restos" sobrantes del proceso de enriquecimiento.

El combustible agotado se compone de materiales de tres clases: productos de fisión, uranio y componente transuránico. Los productos de fisión, que componen aproximadamente el 5 por ciento del combustible usado, son los verdaderos residuos, las cenizas, si se prefiere, de la combustión por fisión. Están constituidos por una mezcolanza de elementos más ligeros creados por escisión de los átomos pesados. Durante sus primeros años, esa mezcla es fuertemente radiactiva. Tras una década, más o menos, en la radioactividad predominan dos isótopos, el cesio 137 y el estroncio 90. Ambos se disuelven en agua, por lo que deben controlarse muy estrechamente. Al cabo de unos tres siglos, la radiactividad de esos isótopos se reduce a la milésima parte, momento en que ya son prácticamente inocuos.

El grueso del combustible nuclear agotado (en torno al 95 por ciento) está formado por uranio, un uranio no fisionado que ha perdido la mayor parte de su isótopo 235 y se asemeja al uranio natural (que en un 0,71 por ciento consiste en uranio 235 fisible). Este componente no es demasiado radiactivo y, si se separa de los productos de fisión y de los demás materiales del combustible agotado, podría almacenarse fácilmente y sin peligro, para su uso futuro, en instalaciones poco protegidas.

El resto del material, la porción de verdad problemática, es el componente transuránico, los elementos más pesados que el uranio. Esta parte del combustible es sobre todo una mezcla de isótopos del plutonio, con una importante presencia de americio. Aunque los transuránicos forman sólo el uno por ciento del combustible agotado, constituyen el peor problema que plantean hoy los residuos nucleares. Su semivida (período que tarda la radiactividad en reducirse a la mitad) asciende a decenas de miles de años, característica que llevó a los organismos reguladores de EE.UU. a exigir que el proyectado depósito de residuos de alta radiactividad en el monte Yucca (en el estado de Nevada) pueda aislar durante más de diez mil años el combustible agotado.

Una estrategia desfasada

Los primeros ingenieros nucleares esperaban que el plutonio del combustible agotado de los reactores térmicos sería extraído y luego empleado en reactores nodriza de neutrones rápidos, que produjeran más plutonio del que consumían. Los pioneros de la energía nuclear imaginaron un sistema económico de la energía que comprendería el comercio libre del plutonio. Pero este puede emplearse en la fabricación de armas. Cuando la técnica nuclear dejó de quedar reservada a las grandes potencias, ese posible uso provocó inquietudes acerca de una proliferación incontrolada de armas atómicas entre otros estados e incluso entre grupos terroristas.

En 1968, el Tratado de No Proliferación Nuclear abordó parcialmente ese problema. Los estados que deseasen sacar provecho de las técnicas de la energía nuclear podían firmar el tratado y comprometerse a no adquirir armas nucleares a cambio de que los países poseedores de armas les ayudasen en las aplicaciones pacíficas. Aunque desde entonces un cuadro de inspectores internacionales ha supervisado el cumplimiento del tratado por sus

Resumen/Reciclado nuclear

- Para reducir al mínimo el calentamiento global, gran parte de la producción futura de energía quizá deba proceder de reactores nucleares, que no liberan dióxido de carbono.
- Pero si se construyeran muchas más centrales nucleares térmicas (es decir, de neutrones lentos, como casi todas las hoy existentes), las reservas mundiales de mineral de uranio barato se agotarían en cuestión de décadas. Además, las grandes cantidades de residuos radiactivos producidos sólo en EE.UU. tendrán que guardarse durante al menos 10.000 años, mucho más de lo que puede albergar el cementerio nuclear del monte Yucca. Peor aún, una fracción importante de la energía que podría extraerse del mineral de uranio original quedaría atrapada en esos residuos.
- Con un nuevo ciclo del combustible de rendimiento mucho mayor, basado en reactores de neutrones rápidos y un reciclado pirometalúrgico del combustible agotado, se aprovecharía en mucho mayor grado la energía contenida en el mineral de uranio extraíble de las minas. Ese ciclo reduciría la creación de residuos de larga vida y mantendría indefinidamente la generación de energía nuclear.

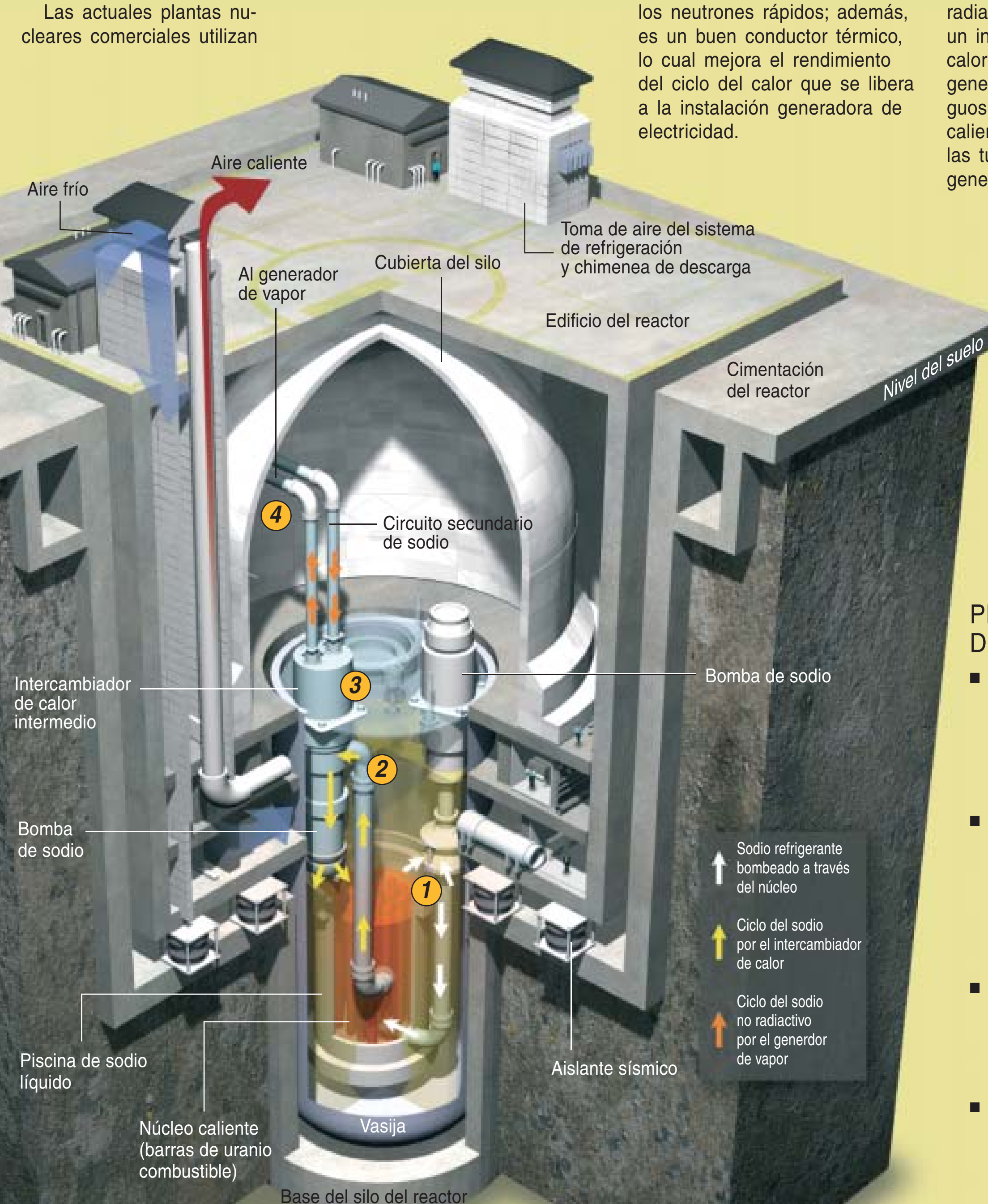
UN NUEVO TIPO DE REACTOR NUCLEAR

Un ciclo de generación de energía menos peligroso y más sostenible podría basarse en los reactores avanzados de metal líquido (RAML); el Laboratorio Nacional de Argonne estudió este sistema en los años ochenta. Un sistema fundado en el RAML emplearía las reacciones en cadena que tienen lugar en el núcleo del reactor como fuente de calor para generar electricidad.

Las actuales plantas nucleares comerciales utilizan reactores térmicos, que hacen uso de neutrones lentos para cebar reacciones en cadena en uranio o plutonio combustibles. En cambio, un sistema basado en un RAML emplearía neutrones rápidos (más energéticos). Consumiría así la totalidad de los núcleos de uranio y elementos más pesados, con lo que se aprovecharía muchísimo más la energía del combustible. En un principio, un reactor nuevo gastaría combustible obtenido del reciclado de combustible agotado procedente de los reactores térmicos.

En la mayoría de los tipos de reactores térmicos, el núcleo está sumergido en agua, que lo mantiene frío y refrena (modera) los neutrones. En un RAML, sin embargo, por refrigerante se emplea una piscina de sodio líquido circulante (*1*). Se eligió este metal porque no modera los neutrones rápidos; además, es un buen conductor térmico, lo cual mejora el rendimiento del ciclo del calor que se libera a la instalación generadora de electricidad.

Un reactor rápido funciona así: el calor nuclear generado en el núcleo calienta el sodio líquido radiactivo que lo atraviesa. Parte de ese sodio caliente se bombea a un intercambiador de calor intermedio (*2*), donde transfiere su energía térmica a un sodio líquido no radiactivo que fluye por los tubos contiguos, aunque separados, de un circuito de sodio secundario (*3*). A su vez, ese sodio no radiactivo (*4*) aporta su calor a un intercambiador/generador de calor (*no se muestra*), donde se genera vapor en tubos contiguos llenos de agua. Ese vapor caliente a gran presión mueve las turbinas que accionan los generadores eléctricos.



DON FOLEY

PROTECCION DEL REACTOR

- Durante el funcionamiento, bombas de gran potencia impulsan el sodio refrigerante a través del núcleo. Si fallaran, el refrigerante circularía por gravedad.
- Para rebajar la temperatura del núcleo, en caso de mal funcionamiento o parada de las bombas, los dispositivos de seguridad expulsarían neutrones del núcleo.
- En las emergencias, seis barras absorbentes de neutrones descenderían hasta el interior del núcleo para pararlo en el acto.
- Si prosiguiera la reacción en cadena, se soltarían en el interior del núcleo miles de bolas de carburo de boro, absorbente de neutrones.

firmantes, la eficacia del acuerdo ha sido desigual porque carece de autoridad efectiva y de medios ejecutivos y coercitivos.

Los ingenieros de armamento nuclear necesitan plutonio con un gran contenido del isótopo 239, mientras que el plutonio procedente de las centrales comerciales suele contener cantidades considerables de los otros isótopos, lo que dificulta su uso militar. Sin embargo, no es inconcebible que se emplee en armas plutonio procedente de combustible agotado. Por ello, en 1977 el presidente Jimmy Carter prohibió en EE.UU. la regeneración de combustibles nucleares. Aducía que, si no se recuperaba el plutonio del combustible nuclear, no podría ser empleado en la fabricación de armas. Deseaba también que Estados Unidos diera ejemplo al resto del mundo. El ejemplo, sin embargo, no lo han seguido Francia, Japón, Rusia y el Reino Unido, por lo que en varios países se sigue reciclando plutonio con destino a centrales eléctricas.

Una alternativa

Cuando se dictó la prohibición, "reciclado" era sinónimo de método purex ("*p*lutonium *ur*anium *ex*traction"), técnica desarrollada para satisfacer las necesidades de plutonio químicamente puro en armas atómicas. Pero los reactores avanzados de neutrones rápidos permiten una estrategia alternativa de regeneración sin que intervenga el plutonio en ninguna de sus fases. Los reactores rápidos pueden así minimizar el riesgo de que el combustible agotado en la generación de energía se aproveche para la fabricación de armas, a la vez que brinda una extraordinaria capacidad para exprimir la máxima energía del combustible nuclear. De esos reactores se han construido varios —en Francia, Japón, Rusia, Reino Unido y EE.UU.—, dos de los cuales aún están en servicio.

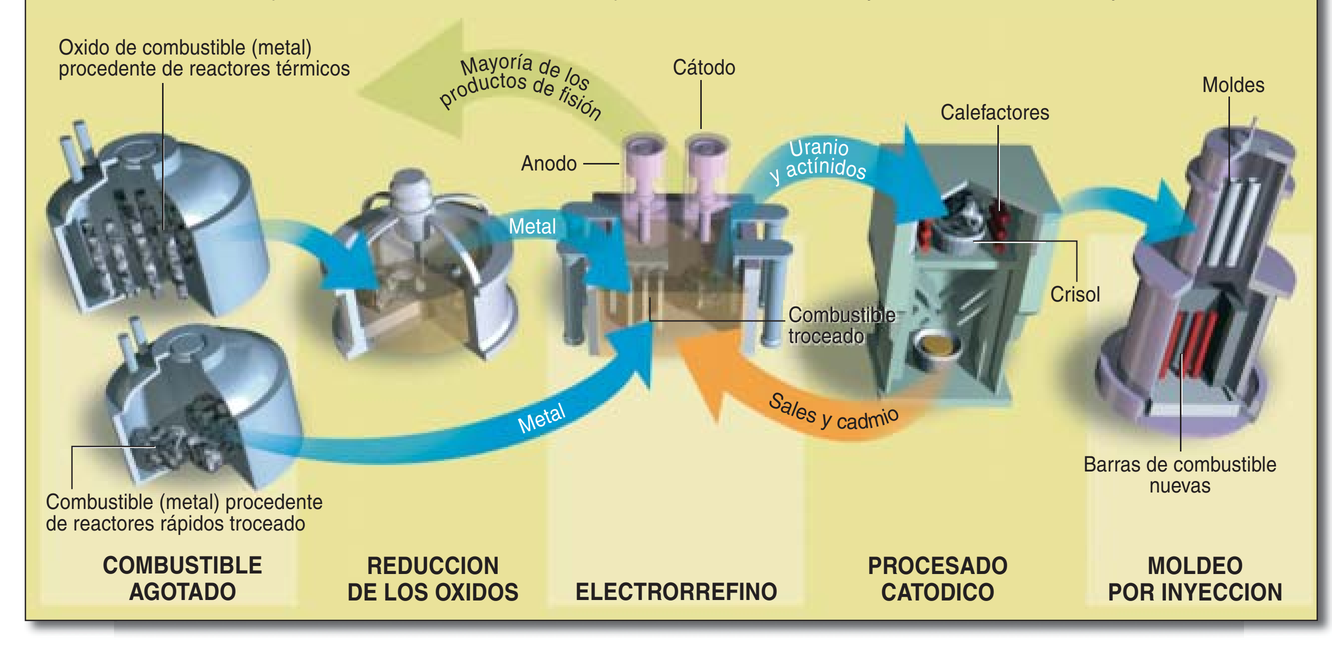
Los reactores rápidos pueden extraer más energía del combustible nuclear porque sus neutrones de gran velocidad (más energéticos) provocan fisiones atómicas más eficientemente que los lentos neutrones térmicos. Esa eficacia tiene su origen en dos fenómenos. En primer lugar, a velocidades bajas, reacciones que no son de fisión absorben muchos más neutrones, que, por lo tanto, se pierden para la generación de energía de fisión. En segundo, la mayor energía de un neutrón rápido aumenta mucho la probabilidad de que un átomo de metal pesado fértil, de uranio 238 por ejemplo, se fisione al ser alcanzado. A causa de ello, en un reactor rápido no sólo es probable que el uranio 235 y el plutonio 239 se fisionen; también lo hará una fracción apreciable de los átomos transuránicos.

En los reactores rápidos no puede emplearse agua para extraer el calor desde el núcleo, porque moderaría los neutrones rápidos. Lo habitual es que se emplee un metal líquido, sodio, en particular, como refrigerante y vehículo térmico. El metal líquido posee una enorme ventaja sobre el agua. Los sistemas refrigerados por agua funcionan a una

UN NUEVO PROCEDIMIENTO PARA REUTILIZAR EL COMBUSTIBLE NUCLEAR

La clave del reciclado pirometalúrgico del combustible nuclear se esconde en el refino electrolítico, o electrorrefino. Separa el verdadero residuo, los productos de fisión; deja aparte uranio, plutonio y demás actínidos (los elementos radiactivos pesados) contenidos en el combustible agotado. Los actínidos permanecen mezclados con el plutonio para que éste no pueda usarse directamente en un arma nuclear.

El combustible agotado procedente de los reactores térmicos hoy en servicio (óxidos de uranio y plutonio) se sometería primero a una reducción para convertirlo en metal, mientras que el uranio y el plutonio procedentes de los reactores rápidos entrarían directamente en la cuba electrolítica. El electrorrefino se asemeja a la galvanoplastia: el combustible agotado, incorporado a un ánodo, se suspendería en un baño químico; a continuación, la corriente eléctrica extraería el uranio y otros actínidos para depositarlos en el cátodo; esos elementos extraídos se enviarían a un procesador catódico para eliminar las sales y el cadmio residuales que dejó el refino; por último, el uranio y los actínidos resultantes se moldearían en barras de combustible nuevo y se reciclarían las sales y el cadmio.



DON FOLEY

presión muy elevada; una pequeña fuga puede trocarse enseguida en un gran derrame de vapor y quizás en una rotura grave de alguna tubería, con una pérdida rápida del refrigerante del reactor. Los sistemas de metal líquido, en cambio, trabajan a la presión atmosférica, por lo que la posibilidad de vertidos importantes es muchísimo menor. Sin embargo, el sodio arde en contacto con el agua, por lo que debe manejarse con cuidado. Con los años, la industria ha acumulado una experiencia considerable en el manejo de esas sustancias y dispone de procedimientos contrastados. Pero ha habido incendios promovidos por el sodio, y habrá más sin duda. En 1995 se desató uno en el reactor rápido de Monju, en Japón. En el edificio del reactor se produjo una gran confusión, pero en ningún momento representó una amenaza para la integridad de la máquina, ni tampoco hubo lesiones ni irradiaciones. Los ingenieros no consideran que la inflamabilidad del sodio constituya un problema grave.

Los investigadores del Laboratorio Nacional de Argonne empezaron a desarrollar la técnica de los reactores rápidos en los años cincuenta. En los ochenta, esa investigación se orientó hacia un reactor avanzado de metal líquido (RAML), de combustible metálico refrigerado por metal líquido, que debía integrarse en una instalación de tratamiento pirometalúrgico para reciclar y reponer el combustible. Se han estudiado otras formas de reactor rápido, algunas con uranio o plutonio metálicos como combustible, otras con combustibles de óxido. Por refrigerantes, se han utilizado plomo líquido o soluciones plomo-bismuto. El combustible metálico que emplea el RAML es preferible al de óxido por varias razones: resulta más seguro, crea combustible nuevo con mayor celeridad y se combina mejor con el reciclado pirometalúrgico.

Pirotratamiento

El tratamiento pirometalúrgico no extrae del combustible usado plutonio puro, como el método PUREX, sino una mezcla de elementos transuránicos. Se basa en la galvanoplastia —el uso de electricidad para recoger, sobre un cátodo conductor metálico, un metal extraído en forma iónica de un baño químico—. Debe su nombre a la alta temperatura a la que debe someterse al metal durante el proceso. Se han desarrollado dos métodos, uno en EE.UU. y otro en Rusia. La diferencia más significativa es que el método ruso trata combustible cerámico (óxido), mientras que en un RAML el combustible es metálico.

En el método norteamericano, se disuelve el combustible metálico agotado en un baño químico. Una intensa corriente eléctrica recoge selectivamente el plutonio y los otros elementos transuránicos en un electrodo, junto con algunos productos de fisión y gran parte del uranio. En el baño permanecen la mayoría de los productos de fisión y algo de uranio. Cuando se ha acumulado un lote completo, se retira el electrodo, del que se desprende mecánicamente el material acumulado. Este se funde y moldea en lingoteras, y los lingotes se reelaboran para convertirlos en combustible de reactor rápido. Cuando el baño se satura con los productos de fisión, se limpia el disolvente y se procesan los productos de fisión extraídos, para su eliminación permanente.

Así pues, a diferencia del actual método PUREX, el método pirometalúrgico recoge virtualmente todos los elementos transuránicos (plutonio incluido), con unos remanentes considerables de uranio y productos de fisión. Sólo una pequeña fracción de los componentes transuránicos va a parar al flujo final de residuos; así se reduce drásticamente el tiempo de aislamiento necesario. La combinación de productos de fisión y transuránicos resulta inapropiada en la fabricación de armas; tampoco vale para un reactor térmico. Pero esa mezcla se torna no sólo aceptable, sino ventajosa a la hora de alimentar reactores rápidos.

Aunque el reciclado pirometalúrgico no está del todo listo para su uso comercial inmediato, ya se ha demostrado la viabilidad de sus principios básicos: en centrales eléctricas en servicio, tanto de EE.UU. como de Rusia, se ha llevado a cabo de manera experimental. Pero aún no ha funcionado a la escala real de la producción.



LABORATORIO NACIONAL DE ARGONNE

EL PROCESAMIENTO PIROMETALURGICO deposita el uranio y los actínidos extraídos del combustible agotado de los reactores térmicos en el cátodo de un baño electrolítico. Tras un tratamiento ulterior, el combustible metálico puede consumirse en reactores de neutrones rápidos.

Comparaciones

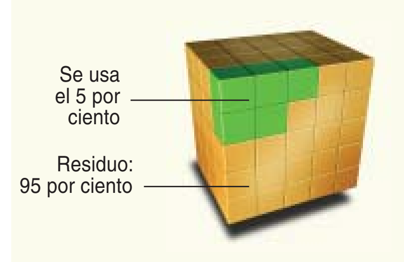
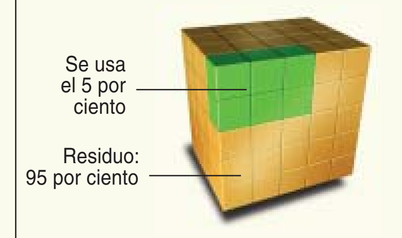
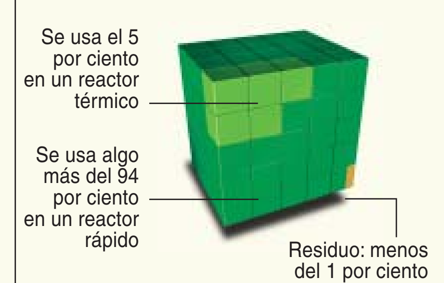
Las características de explotación de los reactores térmicos y rápidos son similares en algunos aspectos; en otros, las diferencias son enormes. El reactor térmico de una planta productora de 1000 megawatt de potencia eléctrica genera al año más de 100 toneladas de combustible desechado. En cambio, la producción anual de residuos de un reactor rápido de la misma capacidad eléctrica es de poco más que una tonelada de productos de fisión, con trazas de transuránicos.

Con el ciclo del RAML, la gestión de los residuos se simplificaría: los residuos de los reactores rápidos no contendrían cantidades grandes de transuránicos de vida larga y sus radiaciones disminuirían al nivel del mineral del que proceden en cuestión de centenares de años, y no de decenas de miles.

Si sólo se emplearan reactores rápidos, el transporte de materiales

COMPARACION DE TRES CICLOS DE COMBUSTIBLE

Hay tres métodos de quemar el combustible nuclear y manejar sus residuos. Aquí se indican sus características principales.

EN UNA SOLA DIRECCION	RECICLADO DE PLUTONIO	RECICLADO COMPLETO
El combustible se quema en reactores térmicos y no se reprocesa; así se hace en Estados Unidos	El combustible se quema en reactores térmicos, tras lo cual se extrae el plutonio mediante el proceso llamado PUREX; así se hace en varios países desarrollados	En los reactores avanzados de neutrones rápidos se quema combustible reciclado por pirometalurgia. De esta técnica sólo se han ensayado prototipos
EMPLEO DEL COMBUSTIBLE		
 Provisión inicial de combustible		
Consume alrededor del 5 por ciento de la energía contenida en el combustible del reactor térmico y menos del 1 por ciento de la contenida en el mineral de uranio (que es la fuente original del combustible) No puede quemar uranio empobrecido (la parte que se retira cuando se enriquece el mineral) ni el uranio del combustible agotado	Aprovecha alrededor del 6 por ciento de la energía contenida en el combustible del reactor térmico y menos del 1 por ciento de la contenida en el mineral de uranio No puede quemar uranio empobrecido ni el uranio del combustible agotado	Puede recuperar más del 99 por ciento de la energía contenida en el combustible agotado del reactor térmico Cuando se acabe el combustible agotado de reactor térmico, puede quemar uranio empobrecido para recuperar más del 99 por ciento de la energía que queda en el mineral de uranio
INSTALACIONES Y OPERACIONES NECESARIAS Rojo: requiere protecciones físicas estrictas • Naranja: sólo necesita unas protecciones físicas modestas • Azul: riesgo potencial para próximas generaciones		
Minas de uranio Enriquecimiento del combustible para concentrar el uranio fisible Fabricación del combustible Centrales eléctricas Un almacén transitorio de residuos (hasta que reciban un destino definitivo) Un depósito permanente capaz de aislar los residuos sin peligro durante 10.000 años (No necesita manipulación del plutonio ni procesado del residuo)	Minas de uranio Enriquecimiento del combustible Mezclado del plutonio Fabricación del combustible en otro lugar Reciclado PUREX en otro lugar Centrales eléctricas Un almacén transitorio de residuos Procesado de los residuos en otro lugar Un depósito permanente capaz de aislar los residuos sin peligro durante 10.000 años	Fabricación *in situ* del combustible Procesado pirometalúrgico *in situ* (reciclado inmediato de residuos) Centrales eléctricas Procesado *in situ* de residuos Un depósito capaz de aislar los residuos durante menos de 500 años (No hará falta extracción minera durante siglos; nunca más habrá que enriquecer uranio)
DESTINO DEL PLUTONIO		
Aumento de contenidos de plutonio en el combustible usado Por mezcla con combustible nuevo, sólo se consigue una degradación lenta del exceso de plutonio de calidad militar	Aumento de contenidos de plutonio en el combustible usado, disponible para uso comercial Por mezcla con combustible nuevo, sólo se consigue una degradación lenta del exceso de plutonio de calidad militar	Los contenidos menguan hasta que sólo quede lo que se emplea en los reactores o se esté reciclando. Los excedentes de plutonio de calidad militar puede degradarse rápidamente El plutonio contenido en el combustible no es utilizable para armas porque resulta demasiado impuro
TIPOS DE RESIDUOS		
Un combustible usado, rico en energía, que se aísla en contenedores almacenados en una instalación subterránea El residuo es tan radiactivo, que durante varios siglos se “protege por sí mismo” de la mayoría de los grupos que quisieran conseguir plutonio 239 para construir armas nucleares	Un residuo vítreo muy estable, rico en energía El residuo es tan radiactivo, que durante varios siglos se “protege por sí mismo” de la mayoría de los grupos que quisieran conseguir plutonio 239 para construir armas nucleares	Un residuo conformado en trozos a la medida que sólo tendrían que mantenerse intactos durante 500 años, tras los cuales el material dejaría de ser peligroso Faltos de plutonio, los residuos no serían útiles para fabricar armas

DON FOLEY

UNA OPINION DIFERENTE

A los reactores nucleares de cuarta generación que se describen en este artículo se los suele considerar una utopía lejana. Sin embargo, Hannum, Marsh y Stanford insisten aquí en la necesidad de poner enseguida en funcionamiento un sistema de reactores nucleares que se base en recursos perdurables y permita una gestión aceptable de los residuos. Es una actitud laudable, pues persigue poner la industria nuclear a la altura del problema energético que habrá que afrontar en los próximos decenios: la sustitución a gran escala de los combustibles fósiles (gas, petróleo, etc.).

El artículo presenta la pirometalurgia como la condición *sine qua non* para que los reactores de neutrones rápidos refrigerados por sodio eviten la proliferación de materiales nucleares y reduzcan la cantidad de elementos radiactivos destinada a acabar como residuo irrecuperable. No está tan claro. Las investigaciones efectuadas en Francia han demostrado que por vía acuosa pueden recuperarse no sólo el uranio y el plutonio, sino también otros actínidos menores. No llega así al 0,1 por ciento la pérdida de actínidos, bastante menos que en la pirometalurgia.

Hay otro punto oscuro: el aporte inicial de plutonio necesario para que arranque un gran número de reactores rápidos. Desde este punto de vista, la pirometalurgia no parece prometedora. Un "reactor europeo rápido" —proyecto industrial en la línea del reactor rápido francés Superfénix, ya clausurado—, que se refrigeraría con sodio líquido, utilizaría óxidos como combustible y procesaría los desechos mediante el procedimiento PUREX, necesitaría para empezar a funcionar cerca de doce toneladas de plutonio por gigawatt de potencia. Esa es la producción de plutonio durante cincuenta años de un reactor actual. Los reactores de metal líquido avanzados que propugnan los autores no presentan ninguna ventaja a este respecto: aunque es verdad que gracias a la pirometalurgia los combustibles agotados se pueden tratar sin esperar a que se enfríen, ese tipo de reactor sigue necesitando grandes cantidades de plutonio para arrancar. Según los estudios realizados en Francia, la sustitución de todos los reactores térmicos por reactores de neutrones rápidos, de empezar en 2040, no podría concluirse sino en 2090. ¿Cómo cabe imaginar que un país menos nuclearizado que Francia —es decir, todos— pueda basar su desarrollo nuclear sólo en los reactores rápidos? Para ello deberían disponer de uranio muy enriquecido (con un porcentaje de hasta el 20 por ciento de uranio fisible), y diseminar los procesos de enriquecimiento, en contradicción con esa lucha contra la proliferación sobre la que insisten los autores.

La pirometalurgia, pues, no puede por sí sola revolucionar la idea —digna de interés— de reactor rápido refrigerado por sodio líquido. En cambio, parece el único tratamiento que pueda vislumbrarse para los reactores de sales fundidas basados en el ciclo regenerador del torio, que exigen también que se extraiga *in situ* una parte de los productos de fisión. Si pasase de los experimentos de laboratorio al estadio industrial, la pirometalurgia permitiría utilizar toda la potencialidad de ese prometedor tipo de reactor, que elimina *de facto* el problema del aporte previo de plutonio y abriría una nueva forma de generar la energía nuclear del futuro.

Sylvain David
Instituto de Física Nuclear de Orsay

radiactivos se llevaría a cabo bajo dos circunstancias: cuando se enviaran productos de fisión al monte Yucca, o a otro depósito de residuos, y cuando se enviara a un nuevo reactor el combustible de arranque. El comercio con plutonio quedaría eliminado.

Hay partidarios de que EE.UU. acometa un programa extensivo de tratamiento PUREX de combustible de reactor que produzca óxidos de uranio y plutonio mezclados para reutilizarlos en los reactores térmicos. Aunque el método de óxidos mixtos (MOX) se emplea ya para inutilizar los excedentes de plutonio de armamento y no sirven para ulteriores usos militares, sería un error desplegar la infraestructura PUREX mucho más extensa que haría falta para procesar el combustible civil. Lo ganado en recursos sería modesto, mientras que subsistiría el problema a largo plazo de los residuos; el esfuerzo sólo retrasaría por breve tiempo la necesidad de unos reactores rápidos eficientes.

El sistema basado en el reactor rápido ofrece una notable versatilidad. Puede ser un consumidor neto o un productor neto de plutonio, o bien podría explotarse de manera que consumo y producción se igualasen. Como productor neto, el sistema proporcionaría combustible de arranque para otros reactores rápidos. En cuanto consumidor neto, gastaría los sobrantes de plutonio y de material para armas. En el equilibrio, el único combustible que necesitaría la planta sería una inyección periódica de uranio empobrecido (uranio del que se ha eliminado buena parte del uranio fisible 235) para reponer los átomos de metal pesado ya fisionados.

Estudios económicos han revelado que esta técnica podría competir comercialmente con las técnicas nucleares ya existentes. El reciclado pirometalúrgico será mucho más barato que el tratamiento PUREX, pero la viabilidad económica del sistema no podrá conocerse mientras no se lo ponga a prueba.

La economía general de una fuente energética depende no sólo de los costos directos, sino también de lo que los economistas llaman "externalidades", los difícilmente cuantificables costos de los efectos externos que resultan de la explotación de esa fuente. Cuando quemamos carbón o petróleo para obtener electricidad, por ejemplo, nuestra sociedad acepta los efectos sanitarios perjudiciales y los costos ecológicos consiguientes. De ese modo, los costos externos vienen a subvencionar la energía producida con combustibles fósiles, sea directamente o mediante los efectos indirectos sobre el conjunto de la sociedad. Aunque difíciles de estimar, las comparaciones económicas que no tienen en cuenta las externalidades no responden a la realidad, sino que resultan engañosas.

Conjunción de tipos de reactores

Si se pusiesen en servicio reactores rápidos, empezarían quemando combustible ya agotado de reactores térmicos, que se habría reciclado por pirometalurgia. Esos residuos, hoy "temporalmente" almacenados *in situ*, se transportarían a plantas de tratamiento, de las que saldrían tres derivados. El primero, muy radiac-

tivo, comprendería la mayoría de los productos de fisión, con trazas inevitables de transuránicos. Se le daría una forma físicamente estable —quizá la de una sustancia de tipo vítreo— y se le enviaría al monte Yucca o a otro cementerio nuclear permanente.

El segundo derivado englobaría los transuránicos, junto con algo de uranio y productos de fisión. Se le convertiría en combustible metálico para reactores rápidos y así se le transferiría en esa forma a reactores RAML.

El tercer derivado, que abarcaría el 92 por ciento del combustible agotado de reactores térmicos, contendría el grueso del uranio, ya en estado empobrecido. Podría almacenarse en lugar seguro para su uso futuro como combustible de reactores rápidos.

A tal situación no puede llegarse de la noche a la mañana. Si empezáramos hoy, el primer reactor rápido entraría en funcionamiento dentro de unos quince años. Ese lapso concuerda bien con el plazo en que se tiene previsto enviar al monte Yucca combustible agotado de reactores térmicos. Pero cabría también reciclarlo en combustible de reactor rápido.

Conforme los actuales reactores térmicos lleguen al final de su vida, podrían sustituirse por reactores rápidos. Si así se hiciese, no se tendría, durante siglos, que explotar las minas de uranio y nunca más habría que enriquecer el uranio. El reciclado del combustible de los reactores rápidos sería tan eficiente, que las actuales existencias de uranio durarían indefinidamente.

China y la India han anunciado que van a construir reactores rápidos. Se sobrentiende que usarán óxido o carburo como combustible, en vez de metal. Se trata de una opción inferior a la óptima; se la habrá elegido, es de suponer, porque el reprocesado PUREX está maduro, mientras que el tratamiento pirometalúrgico aún no se ha aplicado a escala industrial.

Sólo la energía nuclear podrá satisfacer las necesidades de energía a largo plazo de la humanidad sin perjudicar al medio ambiente. Para que la producción de energía nuclear se pueda mantener a gran escala, las existencias de combustible deberán durar mucho tiempo. Se conseguirá si el ciclo de producción de la energía nuclear posee las características que brindan los reactores RAML y la pirometalurgia. Parece el momento de tomar un nuevo rumbo hacia una manera más sensata de generar energía.

Los autores

William H. Hannum, Gerald E. Marsh y **George S. Stanford** son físicos y trabajaron en el desarrollo de reactores rápidos en el Laboratorio Nacional de Argonne. Hannum ha sido jefe de desarrollo de física nuclear e investigación en seguridad de reactores del departamento de energía de EE.UU. Marsh, miembro de la Sociedad Americana de Física, trabajó para el departamento de defensa de EE.UU. como asesor de tecnología y política estratégica nuclear durante las administraciones Reagan, Bush y Clinton. Las investigaciones de Stanford se centraron en la física nuclear experimental, física de reactores y seguridad de reactores rápidos.

Bibliografía complementaria

BREEDER REACTORS: A RENEWABLE ENERGY SOURCE. Bernard L. Cohen en *American Journal of Physics*, vol. 51, n.º 1; enero 1983.

THE TECHNOLOGY OF THE INTEGRAL FAST REACTOR AND ITS ASSOCIATED FUEL CYCLE. Dirigido por W. H. Hannum en *Progress in Nuclear Energy*, Número especial, volumen 31, n.os 1-2; 1997.

INTEGRAL FAST REACTORS. SOURCE OF SAFE, ABUNDANT, NON-POLLUTING POWER. George Stanford en *National Policy Analysis Paper*, n.º 378; diciembre 2001.

LWR RECYCLE: NECESSITY OR IMPEDIMENT? G. S. Stanford en *Proceedings of Global 2003*. Reunión de invierno de la ANS, Nueva Orleans, 16-20 de noviembre, 2003.

S-PRISM FUEL CYCLE STUDY. Allen Dubberly et al. en *Proceedings of ICAPP'03*. Córdoba, 4-7 de mayo, 2003, documento 3144.


PAUL FUSCO *Magnum Photos*

1. UNA MUJER SIN TECHO, en San Francisco.



Pobreza y enfermedad

Nuevos estudios indican que el estrés psicosocial que provoca la pobreza resulta asombrosamente lesivo para la salud

Robert Sapolsky

La vida de Rudolph Virchow, neurocientífico y activista alemán del siglo XIX, estuvo marcada por dos sucesos dramáticos: un brote de fiebres tifoideas en 1847 y las fracasadas revoluciones de 1848. De aquellas experiencias extrajo dos lecciones. Primera, que la propagación de una enfermedad guarda una estrecha relación con la precariedad de las condiciones de vida. Segunda, que quienes disponen del poder cuentan con medios contundentes para subyugar a quienes no tienen ninguno. Virchow las condensó en un célebre epigrama: "Los médicos vienen a ser los abogados de los pobres."

Los médicos (o los expertos en biomedicina en general) pueden considerarse los abogados de los desfavorecidos porque pobreza y enfermedad suelen ir de la mano. Ser pobre entraña malnutrición, unas condiciones de vida insalubres y un sinfín de otros factores que desembocan en enfermedad. Pero no se trata sólo de que los pobres tiendan a enfermar, mientras el resto de la población está sana. Cuando se examina el nivel socioeconómico o NSE (parámetro que valora la renta, la ocupación laboral, la educación y el tipo de vivienda del individuo), se observa que, al ir bajando desde el estrato más pudiente de la sociedad, cada escalón que se desciende en NSE muestra correlación con un estado de salud peor.

Este "gradiente de NSE" se ha corroborado en todas las sociedades de corte "occidental" con relación a enfermedades respiratorias o cardiovasculares, úlceras, trastornos reumatoides, enfermedades psíquicas y cánceres diversos. Las variaciones que describe no son sutiles: al contrastar los peldaños superiores con los inferiores de la escala NSE, se decuplica el riesgo de sufrir algunas enfermedades. En algunos países, al descender por la escala NSE la esperanza de vida se reduce entre cinco y diez años. En los países occidentales, el gradiente más acusado corresponde a los EE.UU., donde los varones blancos más pobres mueren unos diez años antes que los más ricos.

¿A qué se debe esta correlación entre el NSE y el estado de salud? Una situación socioeconómica inferior puede debilitar la salud. Sin embargo, el razonamiento inverso también sirve, pues una salud deficiente puede ser responsable de un bajo nivel socioeconómico; quienes sufren enfermedades crónicas, por ejemplo, pueden ver muy limitada su educación y su productividad laboral, aparte de generar gastos sin cuento.

No obstante, el grueso de los hechos sugiere que la flecha va de la economía a la salud. El NSE constituye, en cierto momento de la vida, un predictor fiable de los parámetros sanitarios futuros. Entre las pruebas que respaldan esta tesis se cuenta un estudio sobre monjas ancianas estadounidenses. Todas ellas habían profesado en su mocedad; durante muchos años habían compartido dieta, cuidados médicos y alojamiento, lo que permitía descontar esas variables de su estilo de vida. Pese a ello, los patrones de enfermedad, la incidencia de demencia y la longevidad mostraban correlación con su nivel socioeconómico en la fecha en que tomaron el hábito, hacia medio siglo.

Hipótesis fallidas

¿Por qué, pues, penetra el NSE en nuestros huesos e influye en nuestra salud? Las respuestas a primera vista lógicas no llegan al fondo de la cuestión. Podríamos argumentar, por ejemplo, que la atención médica resulta menos accesible y de calidad inferior para los pobres. Tal es cierto para EE.UU., donde una fracción notable de pobres carece del seguimiento de un médico de familia: su atención sanitaria se reduce exclusivamente a las visitas a urgencias.

Pero el razonamiento se desmorona si se aplica en países con otra política sanitaria. Veamos, por ejemplo, los estudios Whitehall, realizados por Michael G. Marmot, del Colegio Universitario de Londres, en el curso de los últimos treinta años. El equipo de Marmot ha registrado un conjunto de gradientes de NSE, muy pronunciados, en una población estratificada, la del personal del Servicio Civil (la Administración Pública) británico, que abarca desde el personal laboral de nivel inferior hasta los altos cargos). Los destinados a tareas de mensajería y los conserjes presentan tasas de mortalidad por cardiopatías crónicas muy superiores a los cargos administrativos o ejecutivos. La falta de acceso a los servicios médicos no explica el fenómeno, pues en el Reino Unido, a diferencia de EE.UU., toda la población está cubierta por un servicio de sanidad pública.

Gradientes similares se observan en los países escandinavos u otros donde el régimen de asistencia sanitaria es universal. Las diferencias siguen siendo notables incluso cuando se tiene en cuenta en qué medida los encuestados hacen uso real de los servicios médicos.

De hecho, se observan gradientes de NSE incluso en enfermedades en las que la atención médica resulta irrelevante. Aunque la probabilidad de sufrir diabetes de tipo I (que se manifiesta en la juventud) o artritis reumatoide no cambia en absoluto por muchos controles médicos, análisis de sangre o escáneres que se practiquen, la incidencia de ambas enfermedades es mayor entre los pobres.

Resumen/Estatuto social y salud

- Las personas con nivel socioeconómico (NSE) bajo corren un mayor riesgo de padecer enfermedades y tienen una menor esperanza de vida que las de clase alta. Las explicaciones tradicionales hablaban de un menor acceso de los pobres a la asistencia médica y de estilos de vida más insalubres, como los asociados al tabaquismo o a la obesidad. Pero se trata de razones que no justifican la correlación entre pobreza y mala salud.
- Nuevos estudios indican que las formas de estrés psicosocial asociadas con la pobreza pueden aumentar el riesgo de sufrir un gran número de enfermedades. El estrés crónico que provoca vivir en un barrio marginal y con elevado índice de violencia, aumentaría la susceptibilidad del individuo a la enfermedad cardiovascular, la depresión y la diabetes.
- Otros trabajos han revelado una correlación entre la desigualdad de ingresos y la mala salud en EE.UU. Algunos expertos ponen de relieve que a la pobreza objetiva (*ser* pobre) se añade una pobreza subjetiva (*sentirse* pobre), que provoca mayor estrés todavía en aquellas comunidades con marcadas desigualdades económicas.

Otra explicación "obvia" se centra en los estilos de vida que atentan contra la salud. En las sociedades occidentalizadas, conforme se desciende por la escala del NSE mayor es la tasa de tabaquismo, alcoholismo y obesidad, y peores las condiciones de vida de los individuos, que suelen residir en barrios contaminados, con elevados índices de violencia y hacinamiento. Los pobres tienen mayores dificultades para acceder al agua potable, a una alimentación sana o al uso de instalaciones deportivas, por no hablar de calefacción en invierno y de aire acondicionado en verano.

Parece, pues, evidente que el bajo nivel socioeconómico penetra en los huesos a través del aumento de los riesgos y la reducción de los factores de protección.

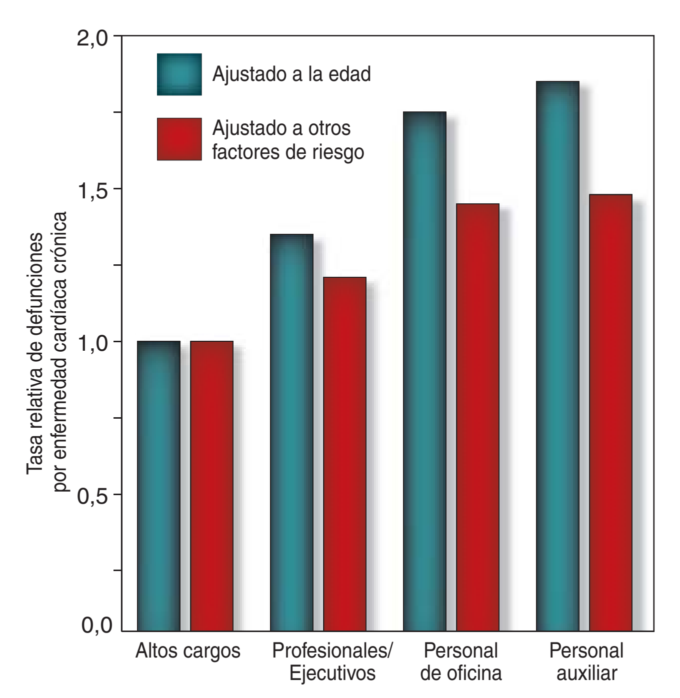
Sin embargo, en la explicación del gradiente de NSE estos factores de riesgo o de protección tienen un escaso protagonismo. En los estudios Whitehall, la influencia del tabaquismo o la práctica de actividad física dieron cuenta sólo de alrededor de un tercio del gradiente. Esta misma tesis se ve refrendada por estudios en los que se comparan riqueza y salud, no dentro de un país, sino en el cotejo entre varios. Parece razonable suponer que cuanto más rico sea un país, de más recursos disponen sus ciudadanos para protegerse y evitar riesgos. Por tanto, la salud debería mejorar progresivamente al ir ascendiendo por el gradiente de riqueza de las naciones, además de hacerlo entre los ciudadanos de cada país concreto. No ocurre así. Entre el 25 por ciento de los estados más ricos del planeta no existe relación entre la riqueza nacional y la salud de sus habitantes.

Así pues, ni el acceso al sistema sanitario, ni el uso que se haga de él, ni la exposición a riesgos, ni la posibilidad de protegerse de ellos explican, en la medida que sería de esperar, el gradiente NSE/salud. Consideraremos, por tanto, otro tipo de factores: las consecuencias psicosociales del NSE.

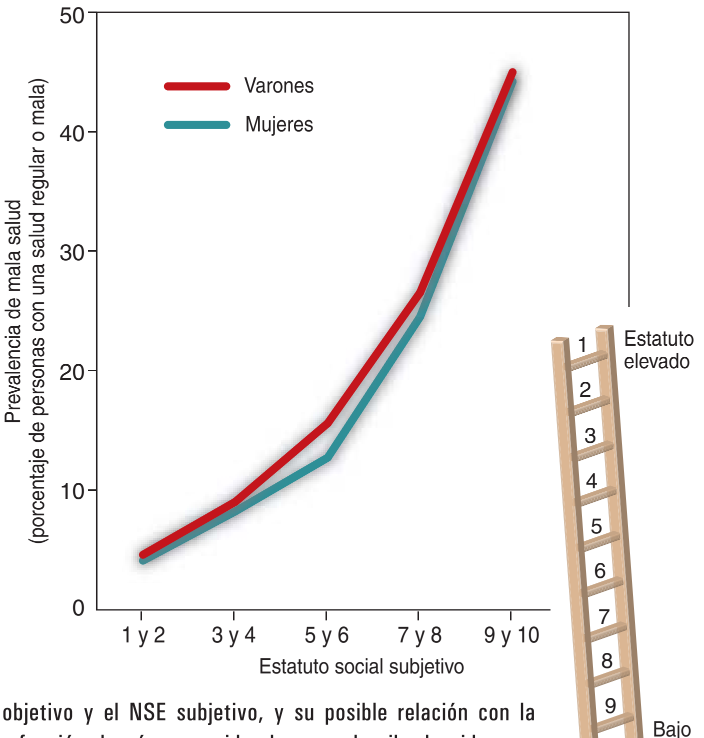
Estrés psicosocial

En una situación ideal, nuestro organismo se halla en equilibrio homeostático, un estado en el que las constantes vitales (ritmo cardíaco, presión arterial, concentración de azúcar en sangre, etcétera) se encuentran en sus intervalos óptimos. Todo aquello que entrañe una amenaza de perturbación de la homeostasis se considera un elemento de tensión. Para la mayoría de los organismos, todo factor de tensión supone un reto físico (piénsese en la necesidad que siente una gacela herida de huir con todas sus fuerzas para salvar su vida, o la que experimenta un depredador hambriento).

Nuestro organismo está excelentemente dotado para afrontar amenazas de corta duración a la homeostasis: se liberan reservas energéticas (glucosa) y el tono cardiovascular aumenta para facilitar la aportación de combustible a los músculos. La digestión, el crecimiento, la reparación de tejidos, la reproducción y otros procesos



2. LOS FUNCIONARIOS BRITANICOS de categoría laboral baja (conserjes y otro personal auxiliar) tienen una probabilidad casi doble de morir por cardiopatías que los altos cargos de la misma edad. Las diferencias en factores de riesgo —mayor incidencia de tabaquismo entre el personal auxiliar, por ejemplo— dan cuenta de menos de la mitad del hiato en las tasas de mortalidad.

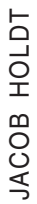


3. EN UN ESTUDIO sobre la diferencia entre el NSE objetivo y el NSE subjetivo, y su posible relación con la salud, se pidió a los participantes que escogieran, en función de cómo consideraban que les iba la vida, una posición en una escalera que simbolizaba la sociedad. Los que declaraban tener una salud frágil o mala abundaban más entre los que se situaron en los peldaños inferiores. El grado notable de correlación entre el estatuto subjetivo y la salud indica que *sentirse* pobre puede incrementar el riesgo de sufrir una enfermedad.

EMILY COOPER; FUENTES: MICHAEL G. MARMOT *Colegio Universitario de Londres* (*gráfica superior*); "SUBJECTIVE SOCIAL STATUS: ITS DETERMINANTS AND ITS ASSOCIATION WITH MEASURES OF ILL-HEALTH IN THE WHITEHALL II STUDY", POR A. SINGH-MANOUX, N. E. ADLER Y M. G. MARMOT, EN *SOCIAL SCIENCE AND MEDICINE*, VOL. 56; 2003 (*gráfica inferior*)



4. EL ESTRES CRONICO puede explicar de qué modo la pobreza repercute en la salud. El riesgo de padecer enfermedades inducidas por estrés aumenta si los individuos carecen de sostén social, no disponen de vías para dar salida a su frustración y perciben que su situación está empeorando (precisamente las condiciones que prevalecen en numerosas comunidades pobres).



fisiológicos que no resultan apremiantes para superar la crisis quedan temporalmente suprimidos. El sistema inmunitario se pone en alerta máxima para neutralizar a patógenos oportunistas. La memoria y los sentidos se agudizan de forma transitoria.

Pero las especies cognitiva y socialmente complejas, los primates como nosotros, experimentamos también otro tipo de estrés. Para los humanos, la mayoría de los factores de tensión implican interacciones con individuos de nuestra propia especie; no suelen perturbar el equilibrio homeostático. Estos elementos psicológicos implican la anticipación (acertada o no) de un reto o problema que habremos de afrontar. La característica más llamativa del estrés psicosocial es la cronicidad: para la mayoría de los mamíferos, el estrés dura sólo algunos minutos; los humanos, en cambio, podemos estar crónicamente angustiados por una hipoteca a 30 años.

Ahora bien, la respuesta de nuestro organismo, aunque adaptativa frente a factores de tensión físicos agudos, resulta patogénica en el caso del estrés psicosocial prolongado. El aumento crónico del tono cardiovascular provoca hipertensión inducida por el estrés. La constante movilización de reservas energéticas acentúa el riesgo o la gravedad de la diabetes de tipo 2 (diabetes de la edad adulta) y otras enfermedades. La inhibición prolongada de la digestión, el crecimiento, la reparación tisular y la reproducción aumenta el riesgo de sufrir diversos trastornos gastrointestinales, frena el crecimiento en los niños, provoca alteraciones de la ovulación en las mujeres y la disfunción eréctil en los varones. Una respuesta inmunitaria al estrés que se prolongue en exceso acaba por suprimir la inmunidad y merma las defensas del organismo. La activación crónica de la respuesta al estrés deteriora la cognición, así como la salud, el funcionamiento e incluso la supervivencia de ciertos tipos de neuronas.

Según la literatura biomédica, los individuos tienen mayor probabilidad de activar una respuesta al estrés y corren mayor riesgo de padecer una enfermedad sensible al estrés si (a) consideran que su capacidad de *control* sobre los estresantes es reducida, (b) sienten que carecen de *información predictiva* sobre la duración e intensidad del factor estresante, (c) disponen de pocas *vías de salida* para la frustración provocada por el estrés, (d) interpretan que el agente estresante es indicativo de un *empeoramiento* de las circunstancias y (e) carecen de *apoyo social* que les ayude a sobrellevar las penalidades provocadas por los elementos de tensión.

Los factores de tensión psicosociales no se hallan distribuidos de forma equitativa en la sociedad. Los pobres no sólo soportan una fracción desmesurada de estresantes físicos (hambre, trabajos que exigen un gran esfuerzo físico, privación crónica de sueño debida al pluriempleo y a camas incómodas que no pueden reemplazar por falta de medios, etcétera) sino que, además, sufren una carga agobiante de factores de tensión psicosociales. El trabajo estupefaciente en la cadena de montaje o una vida laboral ocupada en recibir y acatar órdenes merman y erosionan el sentimiento de autocontrol del trabajador. Una avería en el coche a primera hora de la mañana y las dificultades para estirar la paga hasta fin de mes infligen sensación de impredictibilidad. La pobreza rara vez consiente el acceso a instalaciones deportivas, a la práctica de actividades relajantes (caras) u otros recursos susceptibles de aliviar el estrés, ni años sabáticos que permitan reconsiderar las propias prioridades. Las clases pobres cuentan con menor apoyo social que las clases medias o altas, debido al pluriempleo, los interminables desplazamientos para ir y venir del trabajo y muchos otros incordios.

Marmot ha demostrado que, con independencia del NSE, cuanta menor es nuestra autonomía en el empleo, peor es nuestra salud cardiovascular. El bajo autocontrol en el puesto de trabajo da cuenta de alrededor de la mitad del gradiente NSE para enfermedades cardiovasculares en la población del estudio Whitehall.

Pobreza subjetiva

Otras tres líneas de investigación avalan la influencia del estrés psicológico en los gradientes de salud que guardan relación con el NSE. En el curso de los últimos diez años, Nancy E. Adler, de la Universidad de California en San Francisco, ha estudiado la diferencia entre el NSE objetivo y el NSE subjetivo, y su posible relación con la salud. A los probandos se les presentó un sencillo gráfico: una escalera de mano con 10 peldaños, que simbolizaba la sociedad. Se les pidió entonces que se situaran en uno de los peldaños en función de cómo

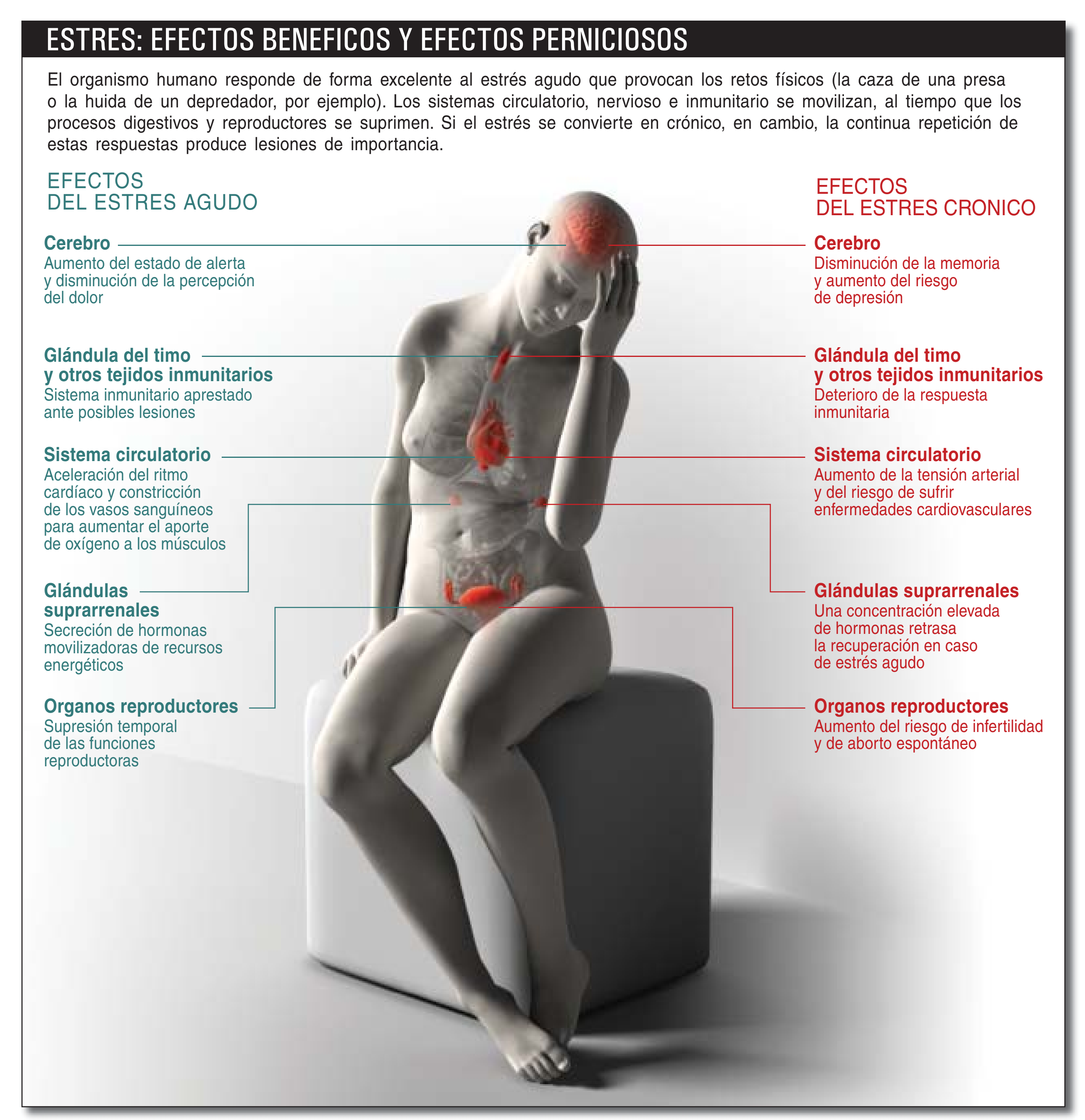
les iba la vida. La pregunta, al ser tan abierta, permitía que el sujeto se identificase con el grupo de referencia por el cual sentía mayor afinidad emocional.

Como Adler ha demostrado, la evaluación subjetiva que hace cada persona de su NSE tiene en cuenta las medidas objetivas habituales (educación, ingresos, empleo, vivienda) así como parámetros relacionados con la satisfacción vital y la preocupación ante el futuro. Lo sorprendente es que el NSE subjetivo resulta cuando menos igual de válido que el NSE objetivo para pronosticar patrones de la función cardiovascular, medidas del metabolismo, incidencia de la obesidad y concentración de hormonas indicadoras de estrés. Las percepciones subjetivas pueden, a buen seguro, contribuir a la explicación de los resultados objetivos.

Esta misma conclusión aflora al comparar los gradientes NSE/salud entre países. Un individuo pobre en EE.UU. puede disponer, objetivamente, de mayores recursos económicos para acceder a cuidados médicos y factores de protección que otro individuo de clase media en un país menos desarrollado y tener, pese a ello, menor esperanza de vida. Según ha puesto de relieve Stephen Bezruchka, de la Universidad de Washington, la renta per cápita en Grecia es aproximadamente la mitad de la estadounidense; sin embargo, la esperanza de vida es mayor en Grecia. Una vez se dispone de los recursos necesarios para mantener un nivel de salud básico (mediante una alimentación y vivienda adecuadas), el nivel absoluto de renta ejerce un efecto escaso sobre la salud. Aunque el trabajo de Adler induce a pensar

ESTRES: EFECTOS BENEFICOS Y EFECTOS PERNICIOSOS

El organismo humano responde de forma excelente al estrés agudo que provocan los retos físicos (la caza de una presa o la huida de un depredador, por ejemplo). Los sistemas circulatorio, nervioso e inmunitario se movilizan, al tiempo que los procesos digestivos y reproductores se suprimen. Si el estrés se convierte en crónico, en cambio, la continua repetición de estas respuestas produce lesiones de importancia.



BRYAN CHRISTIE DESIGN

que *ser* pobre (pobreza objetiva) afecta negativamente a la salud, en el núcleo de ese resultado se encuentra el *sentirse* pobre (pobreza subjetiva).

Hacer que uno se sienta pobre

También Richard Wilkinson, de la Universidad de Nottingham, sostiene que los factores psicosociales median la mayor parte del gradiente NSE/salud. Según los estudios que su grupo de investigación viene realizando desde los últimos quince años, el grado de desigualdad de ingresos en una comunidad demuestra mayor capacidad predictiva que el NSE en toda una batería de medidas de salud. En otras palabras, con independencia de los valores absolutos de la renta personal, cuanto mayor sea la brecha económica que separa los más pobres de los más ricos, peor será la salud media. David H. Abbott, del Centro Nacional de Investigación de Primates en Wisconsin, y el autor, junto con sus respectivos equipos, han observado un fenómeno análogo en animales: en numerosas especies de primates, las estructuras sociales menos igualitarias muestran correlación con niveles basales superiores de una hormona clave para el estrés —indicadora de peor salud— entre los animales socialmente subordinados.

Este hallazgo ha levantado polvareda. Se cuestiona de entrada su generalidad. El trabajo original de Wilkinson proponía que las desigualdades económicas guardaban relación con la salud en un gran número de países y comunidades europeas y estadounidenses. Sin embargo, tal relación resulta válida sólo en el país desarrollado con las rentas más desiguales: Estados Unidos.

A escala de ciudades, así como de estados, la desigualdad de rentas en EE.UU. muestra correlación con las tasas de mortalidad en casi todas las edades. ¿Por qué no se observa este fenómeno en Canadá o Dinamarca? Quizá las diferencias de renta en estos países sean demasiado pequeñas para que se perciba la correlación.

Algunos críticos se han preguntado si la conexión entre las desigualdades económicas y la mala salud no corresponderá a un mero artefacto matemático. La relación entre el NSE y la salud sigue una curva asintótica: la reducción de la esperanza de vida y de otras medidas resulta menos pronunciada cuando se desciende desde el peldaño más elevado de la escala social al inmediatamente inferior que al bajar del penúltimo peldaño al más bajo de la escala. Dado que en una comunidad con grandes desigualdades en la distribución de la renta habrá un número elevado de individuos sumidos en lo más hondo, donde las expectativas de salud son lamentables, la esperanza de vida de esa comunidad será inexorablemente inferior que la de una comunidad igualitaria, por razones que nada tienen que ver con los factores psicosociales.

Wilkinson ha demostrado, sin embargo, que una menor desigualdad en la renta pronostica una mejor salud para los pobres y para los ricos. Ello indica que la asociación entre enfermedad y desigualdad económica no responde a ningún artefacto matemático.

Lo mismo que otros, Wilkinson viene defendiendo que a mayor desigualdad de rentas en una comunidad, mayor estrés psicosocial sufren los pobres. Las desigualdades económicas potencian la jerarquización y dificultan el acceso el sostén social (el sostén mutuo, recíproco y cohesivo se da sólo entre iguales). Además, si a uno le recriminan constantemente su pobreza, es probable que deje de sentirse dueño de su vida, que se agraven las frustraciones que la pobreza entraña y que se intensifique la sensación de que la situación empeora.



5. LAS DESIGUALDADES ECONOMICAS parecen exacerbar el estrés provocado por la pobreza. A medida que se va ensanchando la brecha que separa a ricos y pobres, mengua el apoyo social que reciben los pobres y más intensas se tornan sus frustraciones.

NORBERT MILLAUER *AFP/Getty Images*

Si el estudio de Adler demuestra que el sentimiento de pobreza tiene efectos adversos sobre la salud, los trabajos de Wilkinson sobre las desigualdades económicas sugieren que la forma más segura de conseguir que el que tiene apenas lo justo para vivir se sienta pobre se basa en la continua exaltación de la opulencia de los que viven en la abundancia. En nuestra aldea global, se nos encomian continuamente las riquezas de los potentados y famosos, cuyos recursos ridiculizan los nuestros.

John W. Lynch y George A. Kaplan, de la Universidad de Michigan en Ann Arbor, proponen otro mecanismo de intensificación del estrés psicosocial. Su interpretación "neomaterialista" de la desigualdad de rentas reza como sigue. La asignación de recursos a servicios públicos —mejora del transporte público, universalización de la asistencia médica, etcétera— constituye una forma de aumentar la calidad de vida de la clase media.

Sin embargo, cuanto mayor es la desigualdad de rentas en una sociedad, mayor es la brecha económica que separa la clase media de la rica. Cuanto mayor sea esta distancia, menos ganan los ricos al invertir en bienes públicos. Más se beneficiarían, en cambio, si gastaran el dinero de sus impuestos según sus conveniencias (chófer particular, urbanizaciones selectas con vigilancia privada, agua embotellada, escuelas privadas, seguros médicos privados).

En consecuencia, cuanto más desigual sea la distribución de la renta en una comunidad, más incentivos tendrán los ricos para oponerse a gastos públicos destinados a mejorar la salud de la población. Concretamente en EE.UU., cuanta mayor desigualdad económica exista, tanto mayor poder detentarán los ricos para oponerse a las inversiones públicas. Según Evans, economista ex-

perto en sanidad, este supuesto desembocaría en última instancia en “opulencia privada y miseria pública.”

La “secesión de los ricos” empeora el gradiente NSE/salud de dos formas: mediante el agravamiento de las condiciones de vida de los pobres (lo que se traduce, en parte, en un mayor riesgo para su salud) y la intensificación del estrés psicosocial. Si los factores de tensión psicosociales se entretejen con el sentimiento de ser pobre, potenciado por la exhibición de los opulentos, provocarán mayor estrés todavía cuando los ricos se esfuercen en reducir los bienes y servicios accesibles para los pobres.

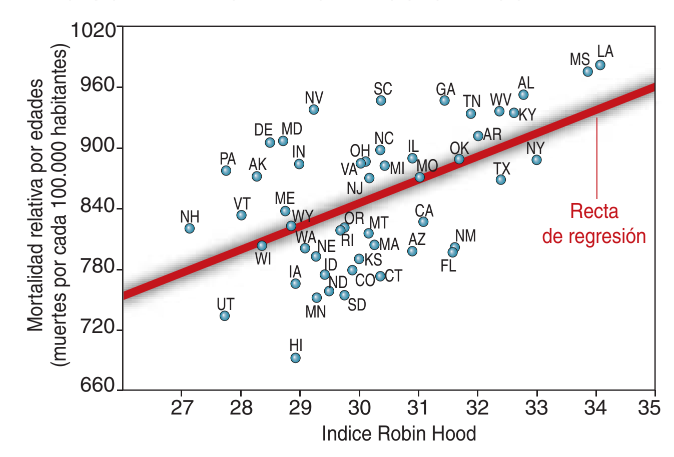
Capital social

Ichiro Kawachi, de la Universidad de Harvard, propone una tercera hipótesis de base psicosocial para explicar la relación entre las desigualdades económicas y la salud. Se apoya en el concepto de “capital social”. Aunque se está precisando su alcance para constituirla en criterio de medida, la noción de capital social se refiere a los niveles de confianza y eficacia que existen en una comunidad. ¿Es norma general que los conciudadanos confíen entre sí y traten de ayudarse? ¿Se sienten motivados los individuos para respetar los recursos que comparten, por ejemplo, conservar los espacios públicos)? ¿Considera la gente que las asociaciones (sindicales, vecinales y otras) sirven para algo? En su mayoría, los estudios de capital social utilizan dos medidas sencillas: cuántas son las organizaciones a las que pertenece cada individuo y la respuesta que reciben preguntas del tipo “¿Cree que la mayoría trataría de aprovecharse de usted, si pudiera?”

Kawachi y otros han mostrado que, en el plano estatal, provincial, ciudadano o de vecindario, un capital social reducido predice una mala “salud objetiva”, esto es, una percepción negativa de la salud personal, junto con elevadas tasas de mortalidad. Mediante el análisis estadístico por caminos (“path analysis”), Kawachi ha demostrado que, una vez descontados los efectos del valor absoluto de los ingresos, la ruta más robusta que conduce de la desigualdad económica a la enfermedad pasa por las mediciones de capital social: es decir, grandes desigualdades en la distribución de rentas vienen acompañadas de escasos grados de confianza y apoyo mutuo, lo que incrementa el estrés y degrada la salud.

Nada de esto resulta sorprendente. La cultura imperante en EE.UU. ha dejado en el abandono a las redes de seguridad social, al tiempo que ha facilitado que los ricos se asienten en lo alto de las pirámides de la desigualdad. Además, se ha optado por renunciar al capital social procedente de las comunidades pequeñas y estables a cambio de oportunidades sin precedente de movilidad y anonimato. Por ello mismo está empeorando la epidemiología social estadounidense. De las naciones “occidentalizadas”, los EE.UU. presentan la distribución de la renta más desigual, con el 40 por ciento de la riqueza controlada por el 1 por ciento de la población; registran también la mayor discrepancia entre el gasto en atención médica (el número 1 del mundo) y la esperanza de vida (puesto 29 en 2003).

De la importancia de los factores psicosociales en la explicación del gradiente NSE/salud se desprende una conclusión crítica: en lo que a salud atañe, la pobreza cuenta mucho más que la falta de dinero. En ocasiones se ha acusado a la escuela psicosocial de promulgar un mensaje antiprogresista: no vale la pena luchar por una cobertura sanitaria para toda la población, ni abaratar los medicamentos y demás medidas de atención médica, porque, a pesar de estas reformas, seguirá existiendo un robusto gradiente NSE/salud. Pero lo que esta investigación enseña no es que hayan de abandonarse tales cambios sociales, sino que deben acometerse muchos más.



6. EL INDICE ROBIN HOOD constituye una medida de la desigualdad económica. Corresponde al tanto por ciento de ingresos totales de la comunidad que debería detraerse de los ricos, cuya renta está por encima de la media, y entregar a los pobres, con una renta inferior a la media, y de ese modo lograr una distribución equitativa. Los estados de EE.UU. (*puntos azules*) con elevado índice Robin Hood propenden también a tener tasas de mortalidad elevadas.

EMILY COOPER; FUENTE: “INCOME DISTRIBUTION AND MORTALITY: CROSS SECTIONAL ECOLOGICAL STUDY OF THE ROBIN HOOD INDEX IN THE UNITED STATES”, POR B. P. KENNEDY, I. KAWACHI Y D. PROTHROW-SITH, EN *BRITISH MEDICAL JOURNAL*, VOL. 312; 1996

El autor

Robert Sapolski imparte clases de biología y neurología en la Universidad de Stanford. Es investigador de los Museos Nacionales de Kenia. Sus trabajos experimentales se centran en los efectos lesivos del estrés sobre el cerebro, así como en terapias génicas para el sistema nervioso. Estudia también poblaciones de babuinos en Africa Oriental para determinar la relación entre la posición social del primate y su salud.

Bibliografía complementaria

MIND THE GAP: HIERARCHIES, HEALTH AND HUMAN EVOLUTION. Richard Wilkinson. Weidenfeld y Nicolson, 2000.

THE HEALTH OF NATIONS: WHY INEQUALITY IS HARMFUL TO YOUR HEALTH. Ichiro Kawachi y Bruce P. Kennedy. New Press, 2002.

THE STATUS SYNDROME. Michael Marmot. Henry Holt and Company, 2004.

WHY ZEBRAS DON’T GET ULCERS: A GUIDE TO STRESS, STRESS-RELATED DISEASES AND COPING. Robert Sapolsky. Tercera edición. Henry Holt and Company, 2004.

Toma de radiocesio por las plantas

El efecto de las inundaciones

La transferencia de un radionucleido del suelo a una planta está controlada por las propiedades del suelo que gobiernan la interacción entre éste y el radionucleido y determinan la cantidad de radionucleido que puede pasar a la planta a través de las raíces (es decir, por absorción radicular). La predicción de la transferencia requiere información específica sobre la interacción: una vez se conocen los mecanismos que controlan la movilidad, es necesario disponer de datos reales para explicar, o incluso predecir, el comportamiento a medio o largo plazo.

El radiocesio ha sido el radionucleido con mayor impacto ambiental tras el accidente de Chernóbil. Años después del accidente, cuando parecía que se había llegado a una situación de estabilidad en la transferencia del cesio del suelo a las plantas, se observó en algunas zonas agrícolas de Bielorrusia y Ucrania un aumento de la cantidad que las plantas encerraban de ese isótopo y de la que se contenía en la leche y la carne, incluso por encima de los niveles a partir de los cuales se prohíbe su consumo. Este incremento guardaría relación con los períodos de inundación en las zonas contaminadas.

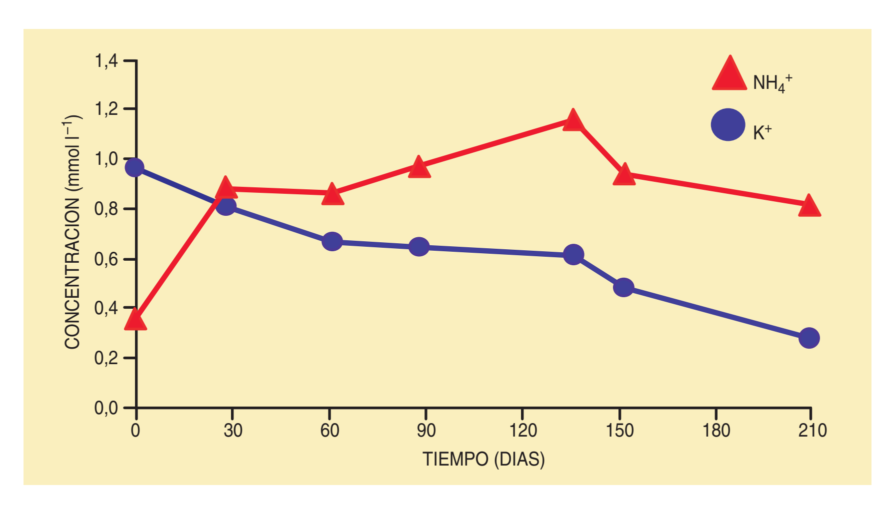
Para estudiar el efecto que tiene la inundación sobre la transferencia suelo-planta, se formularon varias hipótesis sobre los procesos que podían comportar una variación en la toma de radiocesio por las raíces, en el contenido total del radionucleido de las plantas o en ambos.

■ **Variación del contenido de radiocesio en el suelo**. Aunque las plantas absorban la misma proporción relativa de radiocesio, un aumento en el contenido del radionucleido en el suelo comporta un mayor contenido en la planta. En caso de inundación, podrían transportarse pequeñas partículas de suelo enriquecidas en radiocesio (coloides arcillosos) a zonas sin contaminar y permanecer en los suelos tras la retirada de las aguas.

■ **Cambios en la composición de la solución de suelo**. Las dificultades que supone el seguimiento de la solución de suelo ha hecho que se olvide su función clave en el comportamiento de los contaminantes inorgánicos del ecosistema terrestre. (La solución indica el equilibrio entre la fase líquida y la fase sólida de un suelo.) Algunos iones presentes en la solución de suelo, como el potasio o el amonio, compiten con el radiocesio en el proceso de absorción radicular. Por lo general, un aumento de la concentración de amonio comporta un aumento de la toma de radiocesio por la planta, mientras que un aumento de la concentración de potasio causa el efecto contrario. Por lo tanto, cambios en la concentración de estos iones en la solución de suelo, tras una inundación, podrían provocar un cambio en la toma de radiocesio por las plantas.

■ **Modificación de los parámetros de interacción radiocesio-suelo**. El radiocesio presenta una gran afinidad por sitios donde la interacción con el suelo es selectiva. Dicha afinidad depende de la fracción de arcillas de que se trate, en especial, de los filosilicatos. Un cambio en la cantidad de éstos (un incremento por depósito o un decrecimiento por lavado) variará la fracción de radiocesio adsorbida en el suelo, y modificará así la parte que las raíces puedan tomar. Este proceso es tan importante, que se ha definido un parámetro específico, el potencial de intercepción de radiocesio, asociado al número de sitios en un suelo capaces de adsorber selectivamente radiocesio. Así, se puede predecir una disminución de la toma de radiocesio por la planta tras un incremento en la cantidad de filosilicatos en el suelo y un aumento del potencial de intercepción.

Estas hipótesis se examinaron con experimentos a escala, realizados en el laboratorio y en el campo. Como ejemplo de experimentos de laboratorio, la figura muestra los cambios observados en la solución de suelo tras la inundación, durante 210 días, de un suelo con un contenido inicial de potasio mayor que el de la solución que inundaba el suelo. Se observan claramente variaciones en los niveles de amonio y de potasio, que podrían comportar un aumento en la toma de radiocesio por las plantas. Este comportamiento confirmó que, después de una inundación, los niveles de potasio vienen controlados por la composición de la solución inunda-



Cambios en la composición de la solución de suelo, en un terreno con concentración inicial alta de potasio, durante experimentos de inundación.

dora, fuera cual fuese el contenido inicial en el suelo, mientras que los niveles de amonio aumentaban.

Los estudios de campo, aun cuando la variabilidad de sus resultados es mucho mayor que en situaciones controladas de laboratorio, permiten validar las conclusiones de los experimentos de laboratorio y considerar otras variables de difícil estudio en éste. Los resultados confirmaron las conclusiones de los experimentos de laboratorio: aumentaron la solución inundadora control, las concentraciones de potasio y el valor medio de la concentración de amonio. Además, añadieron datos relevantes para la predicción del efecto global de la inundación en la toma de radiocesio por las plantas: se elevaron la concentración de radiocesio, el contenido de filosilicatos y el potencial de intercepción de cesio de los suelos inundados.

Estos estudios llevan a la conclusión de que el proceso de inundación de suelos contaminados podría tener una influencia mayor en la transferencia de radiocesio que la propia dinámica de la interacción de este radionucleido en los suelos afectados. El resultado neto de la inundación, una mayor o menor toma por las plantas, dependerá del balance de las variaciones de todas las propiedades específicas que gobiernan la movilidad de este radionucleido en el sistema suelo-planta.

MARTA CAMPS, MIQUEL VIDAL
y GEMMA RAURET
Universidad de Barcelona

El cofactor de molibdeno

De la biosíntesis a la terapia

El molibdeno (Mo) es un micronutriente esencial para microorganismos, plantas y animales. A excepción de la nitrogenasa bacteriana, el resto de las más de 40 molibdoenzimas descritas tienen un cofactor de molibdeno (Moco, del inglés *Mo*lybdenum *co*factor); es decir, contienen un componente común a todas ellas que incluye molibdeno y participa en la acción de la enzima. En concreto, ese cofactor consiste en una molécula orgánica llamada molibdopterina (MPT), coordinada con un solo átomo de molibdeno. Entre las molibdoenzimas, destacan por su importancia en ciclos clave del metabolismo la nitrato reductasa (NR), fundamental en la asimilación de nitrato; la aldehído oxidasa (AO), que cataliza el último paso en la síntesis de la fitohormona ácido abscísico; la xantina deshidrogenasa (XDH), que interviene en el catabolismo (fragmentación química) de las bases púricas, y la sulfito oxidasa (SO), esencial para la eliminación del sulfito tóxico de la sangre.

Las mutaciones en los genes de la ruta de biosíntesis de Moco provocan la pérdida simultánea de la actividad de todas las molibdoenzimas. La deficiencia del Moco en humanos da lugar a una enfermedad genética autosómica recesiva (es decir, debida a un gen localizado en uno de los cromosomas no sexuales y que sólo se produce cuando la copia paterna y la materna de ese gen presentan la anomalía que causa la enfermedad). Se han identificado numerosos pacientes con deficiencia en Moco que se caracterizan por carecer de SO, AO y XDH. Los pacientes muestran varias anormalidades neurológicas severas (retraso mental, distrofia en cerebro) debido, en parte, a la toxicidad del sulfito y a los bajos niveles de sulfato, anión esencial para la biosíntesis cerebral de esfingolípidos (un tipo de lípidos de las membranas celulares). Hasta el momento no hay terapia disponible; la mayoría de los pacientes muere en la temprana niñez.

Previamente a la biosíntesis de Moco, el molibdeno debe ser transportado al interior celular. Sólo se han identificado los sistemas de transporte de Mo en bacterias; se desconocen los de eucariotas (células dotadas de núcleo diferenciado).

La biosíntesis de Moco se conserva en muy buena medida, desde las bacterias a los humanos; lo mismo ocurre con las proteínas que intervienen en ese proceso. Se pueden distinguir tres etapas básicas (*véase* a *en la figura*). La primera es la conversión de un nucleótido de guanina (GTP) en el precursor Z (PrZ). La segunda es la conversión del PrZ en molibdopterina mediante la adición de dos átomos de azufre (S), reacción catalizada por la MPT sintasa. La tercera es la unión del molibdeno a la MPT. El mecanismo de esta última reacción es, sin duda, el más desconocido de la ruta de biosíntesis del Moco. En las bacte-

Ruta de biosíntesis del Moco (*a*). Se muestra en detalle la reacción catalizada por la proteína Cnx1. Ratones *knockout Mocs1*$^{-/-}$, a la izquierda sin tratar y a la derecha tratados con PrZ (*b*).

CORTESIA DE J. REISS (*fotografía*)

rias, ese paso se halla catalizado por dos proteínas independientes, MoeA y MogA, mientras que en las eucariotas las dos proteínas se encuentran fusionadas en una proteína multifuncional, llamada Cnx1 en las plantas y gefirina en los mamíferos.

La gefirina cumple una segunda función: unir al citoesqueleto los receptores inhibidores de glicina que intervienen en la transmisión del impulso nervioso. Por eso los pacientes con deficiencia en Moco muestran diversos trastornos neurológicos.

Una vez finalizada la síntesis de Moco en las bacterias, se introduce un nucleótido, frecuentemente guanina; se forma entonces el molibdopterín guanín dinucleótido.

Recientemente, se han producido tres grandes avances en el estudio de la estabilidad y mecanismo de biosíntesis de la molécula y en la terapia de la deficiencia en Moco. El primero se refiere a la estabilidad de la molécula y del mecanismo de transferencia, una vez sintetizada, a las desmolibdoenzimas. La vida media de Moco en condiciones aeróbicas es del orden de segundos. Ante ese dato, hubo quien pensó que debía existir algún mecanismo que lo protegiera. El grupo de Emilio Fernández, de la Universidad de Córdoba, ha clonado, purificado y caracterizado la primera proteína que interviene en esa función, concretamente en el alga verde *Chlamydomas reinhardtii*. Se la ha denominado MocoCP (proteína transportadora de Moco). Los autores expresaron heterólogamente la MocoCP —es decir, con otro organismo más apropiado para la investigación, en este caso la bacteria *E. coli*— y determinaron sus características bioquímicas de protección frente al oxígeno y transferencia a las desmolibdoenzimas.

El segundo avance consiste en el esclarecimiento del mecanismo de inserción de Mo en la MPT por el grupo de Ralf Mendel, de la Universidad de Braunschweig. Dicha inserción depende de una activación previa de la molibdopterina mediante una adenilación (la transferencia a una molécula del grupo adenilo del adenosín trifosfato, el transportador de energía a las células) catalizada por la proteína Cnx1. La misma Cnx1 interviene, primero, en la adenilación de la MPT mediante su dominio G (homólogo al MogA bacteriano) y, después, en una desadenilación a través de su dominio E (homólogo al MoeA bacteriano): la presión evolutiva que ha provocado la fusión de las proteínas MogA y MogB de las procariotas. A continuación, al determinar la estructura cristalina de la Cnx1, en el centro activo aparece cobre unido a la MPT. Parece que el cobre intervendría en la protección de los azufres reactivos de la molibdopterina.

El último avance es sin duda el más esperanzador. Se trata de la primera vez que se inyecta PrZ como terapia de ratones *knockout* $Mocs^{1-/-}$ que sirven de modelos de la enfermedad por deficiencia de Moco (*knockout* quiere decir que se les desactiva un gen, en este caso el que rige el primer paso de la biosíntesis de Moco). El proyecto se ha realizado gracias a la colaboración entre el grupo de Jochen Reiss, de la Universidad de Göttigen, que obtuvo los ratones, y el grupo de Mendel, que consiguió purificar el PrZ de un mutante de *E. coli* en elevadas cantidades. Esos ratones knockout $Mocs1^{-/-}$ tratados con PrZ pueden sintetizar de nuevo Moco. Los ratones transgénicos no muestran diferencias físicas ni de comportamiento respecto a los ratones silvestres, y alcanzan la madurez física y sexual, mientras que los no tratados mueren al poco de nacer. Se abre, pues, una ventana esperanzadora para el futuro tratamiento de pacientes humanos afectados por esta enfermedad letal.

Angel Llamas Azúa
y José Angel Santamaría Araujo
Instituto Botánico,
Universidad Técnica de Braunschweig

Genética forense

Los perfiles de ADN

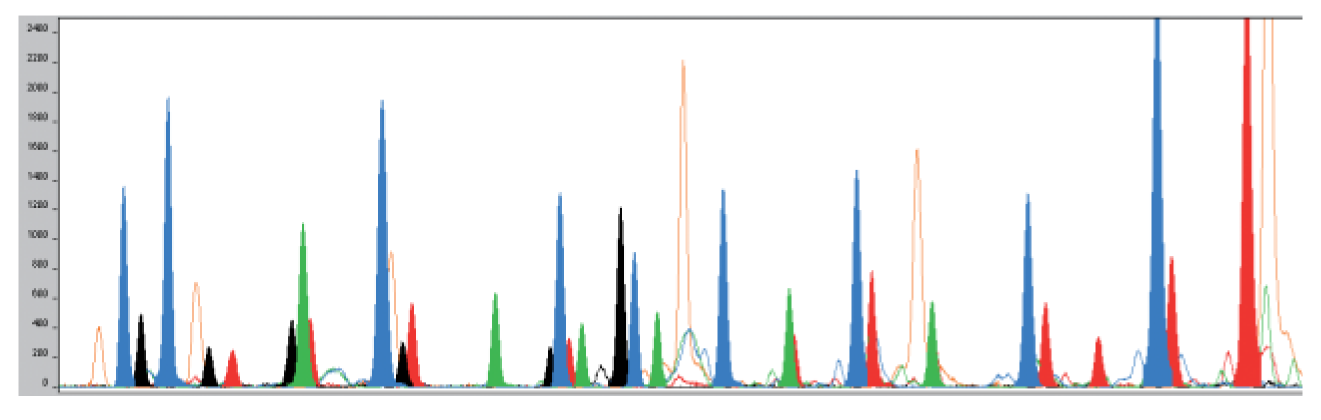
La genética forense, especialidad firmemente implantada en los países desarrollados, consiste en la aplicación de la genética a la resolución de problemas judiciales. Los tribunales recurren a este tipo de peritaje para investigar la paternidad por medios biológicos, estudiar vestigios biológicos de interés criminal —manchas de sangre, esperma, pelos, etc.— o identificar individuos a partir de restos biológicos.

Alec Jeffreys y sus colaboradores describieron en 1985 los perfiles de ADN, conocidos popularmente como huella genética o "prueba de ADN". Desde entonces, su impacto en las ciencias forenses ha sido enorme.

Antes de su descubrimiento, se utilizaban polimorfismos proteicos, analizados mediante técnicas electroforéticas e inmunológicas (la electroforesis separa eléctricamente macromoléculas orgánicas ionizadas en solución según su peso y carga eléctrica). Estos marcadores poseían un valor limitado, particularmente cuando se trataba de material degradado o escaso, como ocurre en la mayoría de los problemas forenses.

Los perfiles de ADN analizan la variación del ADN entre individuos, normalmente de las partes repetitivas que existen en el ADN. Tiene grandes ventajas sobre el análisis de polimorfismos proteicos: aportan mayor información y bastan unas pocas células del material degradado para confeccionarlos —el ADN es mucho más resistente a la degradación que las proteínas—. Además, el mismo genotipo se puede obtener de cualquier tejido, mientras que el análisis de variantes proteicas está restringido a las células donde esos polimorfismos se expresan.

Desde su descubrimiento, las técnicas de análisis y el tipo de polimorfismos de ADN que se estudian experimentaron una continua evolución. En un principio se utilizaban minisatélites de ADN (esto es, repeticiones en tándem de unos 25 pares de bases) muy polimórficos, que se analizaban por medio del método de hibridación Southern (blotting), que consiste en separar los fragmentos de ADN mediante electroforesis en un gel y transferirlos a un soporte de celulosa, donde se les aplica una sonda de ADN que debe hibridarse con los fragmentos para identificarlos. Para ello se empleaban sondas *multi-loci*, que permitían el análisis al mismo tiempo de varios *loci* de minisatélites: sitios del ADN de esos minisatélites. Por motivos de estandarización y problemas en la valoración de la prueba,



Perfil de polimorfismos nucleótidos simples obtenido en un caso criminal: 25 de esos polimorfismos se copian a la vez múltiples veces a partir de una muestra de unos pocos picogramos de ADN, se analizan mediante minisecuenciación y se detectan por electroforesis capilar.

se prescindió de las sondas *multi-loci* y se adoptó el análisis de minisatélites con sondas de *locus* único. Este método, muy laborioso, requería mucho ADN, que tenía además que estar bastante íntegro.

La invención de la reacción en cadena de la polimerasa, procedimiento que se vale de la enzima polimerasa del ADN para crear múltiples copias de un fragmento de ADN —o, como se suele decir, para "amplificarlo"— sin recurrir a la mediación de seres vivos, y el descubrimiento de los polimorfismos de las "repeticiones cortas en tándem", repeticiones de 2 a 6 pares de bases en el ADN, permitieron analizar ADN obtenido de unas pocas células, así como de muestras o restos degradados. La identificación de los restos de los Romanov demostró las posibilidades de esta técnica.

El potencial del ADN ha hecho posible también la resolución de problemas de inmigración o de casos complejos de paternidad en los que el presunto padre ha fallecido o se halla ausente. La identificación rápida de individuos en grandes desastres ha sido también posible gracias a estas técnicas. En la mayoría de los países desarrollados hay bases de datos computerizadas con perfiles de ADN para la identificación de personas que han cometido delitos.

Gracias a los avances en la técnica de fluorocromos y en electroforesis capilar, se utilizan multiplexos (reacciones en cadena de la polimerasa simultáneas de varios fragmentos), que suelen incluir 15 repeticiones cortas en tándem, y el gen de la proteína amelogenina, que sirve para determinar el sexo. Este procedimiento ofrece un enorme poder de discriminación: la probabilidad de que dos individuos no emparentados presenten el mismo perfil es remota; permite, además, el análisis de una muestra de sangre del tamaño de la cabeza de un alfiler en unas pocas horas.

Las repeticiones cortas en tándem que se utilizan, se validan previamente; en general, son repeticiones de cuatro pares de bases. Cuando las muestras están degradadas, se suelen usar microsatélites de pequeño tamaño en el fragmento amplificado por la reacción en cadena de la polimerasa; para el caso de manchas a las que han contribuido dos o más individuos, son convenientes los microsatélites de cinco pares de bases en la unidad de repetición.

El análisis de microsatélites de cromosoma Y es muy importante para el estudio de muestras biológicas donde se mezclan procedencias masculinas y femeninas; por ello, resulta decisivo en delitos sexuales, aunque también se aplica en casos complejos de paternidad.

El análisis por secuenciación —la determinación de sus unidades básicas— de la región hipervariable I y II del bucle-D del ADN mitocondrial resulta también fundamental para el análisis de muestras muy degradadas; en particular, para el análisis de cabellos, ya que habitualmente no contienen raíz o ésta se encuentra en fase telogénica —de crecimiento detenido—, por lo que carece de ADN nuclear.

El futuro del análisis de ADN se halla en los polimorfismos nucleotídicos simples. Todavía se encuentran en fase de investigación y validación. Poseen menos poder de discriminación que las repeticiones cortas en tándem (hacen falta unos 60 polimorfismos nucleótidos simples por cada 15 pares de bases); a cambio, cabe analizarlos con técnicas de alto rendimiento. Entre tales técnicas se cuentan diversos métodos de detección: minisecuenciación, método que detecta mutaciones puntuales o variaciones alélicas mediante la adición específica de nucleótidos a un fragmento de ADN; sondas o sistemas basados en OLA, método que se basa en la capacidad de la ligasa para unir covalentemente dos oligonucleótidos que hibridan uno junto a otro sobre un molde de ADN. Estos métodos pueden ser aplicados en el análisis de variaciones nucleotídicas simples mediante electroforesis, espectrometría de masas o chips de ADN.

Paralelamente a la revolución del ADN en genética forense han tenido lugar dos avances. El primero es la estandarización e introducción de sistemas de control de calidad y acreditación de laboratorios. La Sociedad Internacional de Genética Forense está desempeñando un papel decisivo en la formulación de estándares mundiales. El segundo ha sido la introducción en la valoración de la prueba de la lógica bayesiana, que parte de asignaciones previas de probabilidad, moduladas con la información adicional de que se disponga mediante el teorema de Bayes para determinar las probabilidades condicionales. Esta herramienta permite una valoración más justa de la prueba.

ANTONIO SALAS, MARÍA BRIÓN
y ANGEL CARRACEDO
Instituto de Medicina Legal
Facultad de Medicina
Universidad de Santiago de Compostela

Christiane Hasemann y Covadonga Orejas

Nemátodos del mar profundo



1. Superficie del fondo marino a 2000 metros de profundidad en el estrecho de Fram (Artico). Este tipo de comunidad abisal puede albergar hasta un millón de individuos por metro cuadrado, pertenecientes en muchos casos a especies sin describir.

KAREN VON JUTERZENKA, AWI

El mar profundo constituye la región más enigmática de la Tierra. Conocemos mejor la Luna que las profundidades de los océanos. Más del 53 por ciento de la superficie terrestre se halla cubierta por suelos oceánicos, cuya profundidad oscila entre los 3000 y los 6000 metros.

Noche eterna (a partir de 300 metros no queda rastro alguno de luz solar), presión elevada (hasta una tonelada por centímetro cuadrado), temperaturas bajas (en torno a los 2 grados centígrados) y extrema escasez de fuentes de alimento conforman el medio de las profundidades oceánicas. Pese a su rigor, estas condiciones presentan una ventaja para la vida: la estabilidad.

En el siglo XIX, numerosos investigadores postularon la ausencia total de fauna por debajo de los 500 metros de profundidad. Descubrimientos posteriores demostraron que las aguas oscuras y profundas de los océanos rebosan de vida, mostrando un amplio abanico de formas y estrategias de supervivencia. La oscuridad y la presión elevada que allí reinan han dado lugar a una especialización de los organismos difícil de encontrar en zonas de aguas someras. Se estima que quedan todavía por descubrir más de 10 millones de especies en las profundidades oceánicas. Son varias las hipótesis que explican esta elevada e inesperada biodiversidad; sin embargo, todas convergen en que el sedimento de estas zonas opera como un "microecosistema" marcado por un mosaico de alteraciones biogénicas, que determinan la formación de nichos ecológicos. Esta estructura favorece que el número de especies que colonizan el hábitat sea elevado.

Los nemátodos de vida libre (gusanos filiformes) constituyen el grupo más abundante y diverso de metazoos que medran en el sedimento marino. Representan en muchas zonas del mar profundo más del 90 por ciento de los habitantes del sedimento. Su tamaño oscila entre 42 y 500 micras (meiofauna); con frecuencia se encuentran en concentraciones elevadas (gran cantidad de especies de nemátodos, a menudo del mismo género, en pequeños volúmenes de sedimento).

2. Las diferencias morfológicas entre especies de nemátodos resultan casi imperceptibles. Estos organismos se distinguen sobre todo por su forma de alimentación; ello evita que deban competir entre sí por los recursos tróficos. (Macho del género *Acantholaimus*.)



3. La diversidad trófica junto con las distintas adaptaciones fisiológicas, bioquímicas y etológicas han facilitado la coexistencia entre especies de nemátodos. (Juvenil del género *Daptonema*.)



4. Zona cefálica en la que se observa el aparato bucal. (Género *Sphaerolaimus*.)

CHRISTIANE HASEMANN, AWI

Evolución de los tetrápodos

Ciertos descubrimientos fósiles recientes arrojan luz sobre la evolución de los tetrápodos terrestres a partir de los peces

Jennifer A. Clack

A lo largo de los cuatro mil millones de años transcurridos aproximadamente desde que la vida surgió sobre la Tierra, la evolución ha generado metamorfosis maravillosas. Una de las más espectaculares corresponde a la que, a partir de peces con aletas, dio lugar a los Tetrápodos, animales terrestres con patas y dedos. Este grupo abarca anfibios, reptiles, aves (con sus antepasados dinosaurios) y mamíferos. Algunas especies de los taxones mencionados han modificado o perdido sus extremidades, pero su antecesor común las poseía: dos delante y dos detrás, donde antaño se agitaban las aletas.

La sustitución de aletas por patas constituyó un paso crucial en esta transformación. Pero no fue el único. Cuando los tetrápodos se aventuraron en tierra firme, se enfrentaron a problemas con los que ningún vertebrado se había encontrado hasta entonces. No se trataba sólo de desarrollar patas y comenzar a andar. La tierra constituye un medio radicalmente distinto del agua; para medrar en ella, los tetrápodos tuvieron que desarrollar nuevas formas de respiración, audición, lucha contra la gravedad y un largo etcétera. Tan drástica transformación, sin embargo, les permitió conquistar un nuevo mundo.

Hasta hace unos 15 años, se conocía muy poco acerca de la secuencia de acontecimientos que definieron la transición de peces a tetrápodos. A finales del siglo XIX, Edward D. Cope propuso que los tetrápodos procedían de peces con aletas carnosas, semejantes a las de los peces pulmonados y los celacantos. Pero los detalles de tan fundamental cambio permanecían ocultos. Además, las estimaciones de cuándo sucedió variaban enormemente: desde hace 400 millones hasta hace 350 millones de años, durante el Devónico. El problema radicaba en la parquedad del registro fósil, limitado a un solo pez de este tipo, *Eusthenopteron*, y un tetrápodo devónico, *Ichthyostega*, demasiado avanzado para dilucidar las raíces de su grupo.

Ante tamaña sequía de indicios sobre la naturaleza de la transición, sólo cabía especular. Se hizo famosa entre las demás la conjetura adelantada por Alfred Sherwood Romer, de la Universidad de Harvard. En los años cincuenta del siglo pasado, propuso que peces como *Eusthenopteron*, varados en condiciones áridas, utilizaron sus apéndices musculares para arrastrarse hasta una nueva masa de agua. Se seleccionarían las especies dotadas de capacidad para recorrer una extensión mayor y, por ende, de alcanzar puntos de agua cada vez más alejados. Andando el tiempo, ello habría conducido al origen de extremidades genuinas. Según esta hipótesis, los peces salieron del agua antes de adquirir, por evolución, las patas.

Pero desde entonces han ido descubriéndose numerosos fósiles que aportan datos sobre dicha transformación. Los hallazgos en cuestión han incrementado de forma exponencial nuestro conocimiento sobre ese capítulo clave de la historia de la vida sobre la Tierra, al tiempo que han puesto en cuarentena viejas ideas en torno a la evolución, la diversidad, la biogeografía y la paleoecología de los tetrápodos.

1. RESPIRACION AEREA: *Acanthostega*, un tetrápodo primitivo, saca la cabeza del agua en un pantano situado en lo que hoy es Groenlandia oriental, hace unos 360 millones de años. Aunque poseía cuatro patas, este animal no habría sido capaz de soportar su cuerpo en tierra. Así, en lugar de desarrollar estos miembros para adaptarse a la vida terrestre, parece que inicialmente los utilizó para elevar la cabeza y respirar fuera de un agua pobre en oxígeno. Sólo más tarde encontrarían uso en tierra.

RAUL MARTIN

Desarrollo de las patas

A la primera generación de restos fósiles que configuraron la concepción moderna del origen de los tetrápodos pertenece *Acanthostega*, un animal que vivió



hace unos 360 millones de años en lo que hoy es Groenlandia oriental. En 1952, Erik Jarvik, del Museo Sueco de Historia Natural en Estocolmo, lo identificó sobre la base de dos bóvedas craneales parciales. En 1987, el autor y su equipo hallaron especímenes que mostraban el esqueleto poscraneal de *Acanthostega*.

En numerosos aspectos *Acanthostega* correspondía al tipo intermedio anatómico entre peces y tetrápodos. Sin embargo, de su estudio se extraía una historia distinta de la imaginada: aunque poseía patas y pies, estaba mal equipado para una vida terrestre. Sus patas carecían de tobillos adecuados para soportar el peso del animal en tierra; guardaban mayor semejanza con canaletes para nadar. Si bien poseía pulmones, sus costillas eran demasiado cortas para evitar el hundimiento de la cavidad torácica una vez se hallara fuera del agua. De hecho, muchos de los rasgos de *Acanthostega* eran indiscutiblemente pisciformes: las proporciones de los huesos del antebrazo recordaban las de la aleta pectoral de *Eusthenopteron*; la parte posterior del esqueleto presentaba una cola profunda y en remo con largos radios óseos que debieron de proporcionar el andamiaje para una aleta; amén de pulmones, poseía branquias.

Ese cuadro pisciforme sugería que las extremidades de *Acanthostega* no sólo estaban adaptadas para su uso acuático, sino que ello constituía la condición ancestral de los tetrápodos. Es decir, este animal, aun siendo claramente un tetrápodo, correspondía a un organismo acuático cuyos antepasados inmediatos eran peces que nunca habían abandonado el agua. El descubrimiento obligó a repensar la secuencia de cambios clave para el esqueleto. En vez de presentarnos a un animal como *Eusthenopteron*, que salía del agua arrastrándose y después desarrollaba patas y pies, de acuerdo con la hipótesis de Romer, los nuevos fósiles indicaban que los tetrápodos desarrollaron estos rasgos cuando todavía eran acuáticos; sólo más tarde los emplearían para andar.

Eso significaba, asimismo, que había que volver a cuestionarse las circunstancias ecológicas que ampararon el desarrollo de las extremidades. *Acanthostega* sugería que las exigencias del medio terrestre quizá no constituyeran la fuerza motriz en la evolución temprana de los tetrápodos.

Acanthostega constituía el mejor candidato para erigirse en el eslabón perdido entre los vertebrados terrestres y sus precursores acuáticos. Sin embargo, portaba un rasgo que no correspondía ni a tetrápodo ni a pez: cada una de sus extremidades terminaba en un pie con ocho dedos bien formados, en lugar de los cinco habituales.

Se trataba de un fenómeno desconcertante. Hasta entonces se venía dando por cierto que, en la transición de pez a tetrápodo, el pie pentadáctilo derivó directamente de los dedos que constituían la aleta de *Eusthenopteron* o de un animal parecido. Lo habitual hubiera sido descartar este espécimen, por aberrante. Pero resultó que un esqueleto parcial de *Tulerpeton*, un tetrápodo primitivo hallado en Rusia, poseía un pie hexadáctilo. *Ichthyostega*, otro de los especímenes desenterrados en Groenlandia oriental, poseía también un pie con más de cinco dedos.

Los avances en biología del desarrollo han contribuido a resolver parte de este misterio. Sabemos que hay varios genes, incluidos *Sonic Hedgehog* y la serie *Hox*, que controlan el desarrollo de aletas y extremidades. Los mismos conjuntos de genes se encuentran en peces y tetrápodos, aunque cumplen cometidos diferentes. Así, *Hoxd 11* y *Hoxd 13* desempeñan, tal parece, una función de mayor importancia en los tetrápodos, cuyos dominios de expresión en las yemas de las extremidades son mayores y asimétricos en relación con los de las aletas de los peces. En estas regiones se forman los dedos.

Se desconoce el mecanismo en cuya virtud el pie de cinco dedos surgió del pie de ocho. Sí disponemos de una explicación plausible del motivo por el que el pie pentadáctilo se convirtió en la estructura básica de los tetrápodos: coadyuvaría a la formación de articulaciones del tobillo que ofrecieran, a un tiempo, estabilidad para soportar peso y flexibilidad y, de ese modo, realizar el tipo de ambulación que los tetrápodos acabaron por inventar.

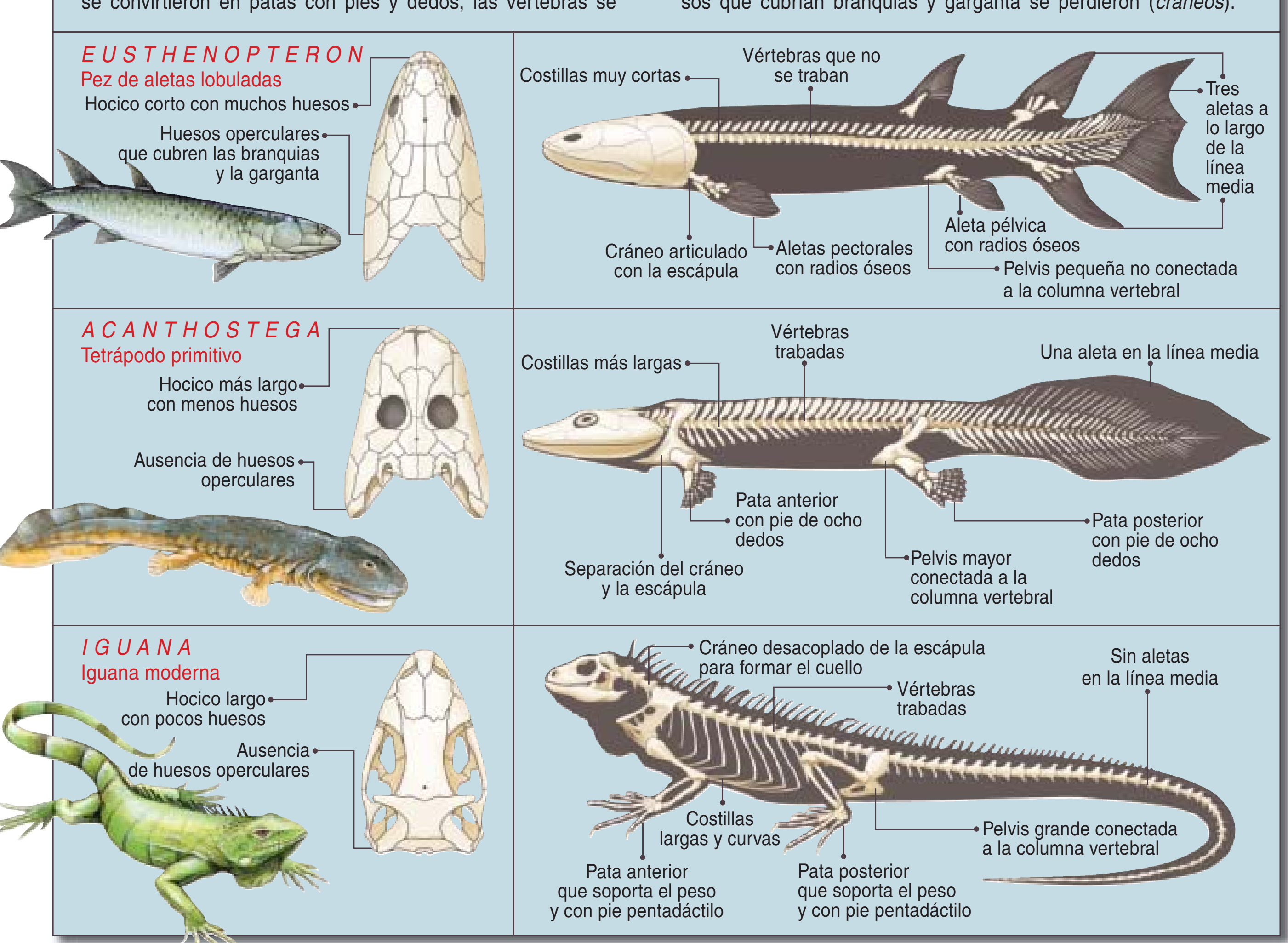
En *Acanthostega* la atención se centró también en el interior de su mandíbula inferior, una parte de la anatomía de los primeros tetrápodos hasta entonces poco apreciada. Por lo general, los peces poseen dos hileras de dientes a lo largo de su mandíbula inferior, con un número elevado de dientes pequeños en la hilera exterior que complementan con un par de colmillos de gran tamaño y algunos dientes pequeños en la hilera interior.

El fósil de *Acanthostega* demostraba que los tetrápodos primitivos poseían una estructura dentaria distinta: un número limitado de dientes mayores en la hilera exterior y una reducción en el tamaño de los dientes alojados en la hilera interior. Tales cambios probablemente acompañaron la transición desde una alimentación exclusivamente acuática hasta la realizada en tierra firme o con la cabeza por encima del agua.

Este hallazgo propició el reconocimiento de otros tetrápodos entre los restos que desde hacía tiempo se amontonaban sin identificar en los depósitos de los museos. Uno de los

Resumen/El origen de los tetrápodos

- La aparición de vertebrados terrestres constituyó un acontecimiento crucial en la evolución de la vida sobre la Tierra.
- Durante decenios, un registro fósil escaso dificultó la investigación de los estadios que condujeron a estos tetrápodos terrestres desde sus antepasados, los peces.
- Fósiles descubiertos en el curso de los últimos 15 años han colmado numerosos vacíos de esa historia evolutiva. Y han arrojado nueva luz sobre la evolución, la diversidad, la biogeografía y la paleoecología de los tetrápodos.
- Los tetrápodos habrían desarrollado numerosos rasgos característicos mientras persistían en su medio acuático. Los miembros primitivos del grupo estaban más especializados y su distribución, geográfica y ecológica, era más amplia de lo que antes se pensaba.



más espectaculares fue el del lituano *Ventastega*, un género del Devónico tardío. En el decenio de 1990, tras el descubrimiento de *Acanthostega*, los expertos repararon en una mandíbula inferior, descubierta en 1933: pertenecía a un tetrápodo. Las excavaciones subsiguientes en el yacimiento donde se encontró *Ventastega* no tardaron en proporcionar más fósiles de una calidad excepcional, incluido un cráneo casi completo.

En el ínterin, han venido apareciendo varios peces cercanos a los tetrápodos que colman la laguna morfológica entre *Eusthenopteron* y *Acanthostega*. De los géneros en cuestión recuperados, dos se conocían desde hacía varios decenios, aunque su estudio detenido se postergó hasta fecha reciente. Nos referimos a *Panderichthys*, un pez poderoso, de hace entre 380 y 375 millones de años de antigüedad, con el hocico puntiagudo y los ojos sobre la cabeza, hallado en la región báltica, y al canadiense *Elpistostege*, muy similar en tamaño y forma a *Panderichthys*, pero de hace de 375 a 370 millones de años de antigüedad.

Elpistostege y *Panderichthys* se encuentran mucho más cerca de los tetrápodos que *Eusthenopteron*. En 2004, una expedición a la isla de Ellesmere, en el Artico canadiense, dirigida por Neil Shubin, de la Universidad de Chicago, recolectó algunos restos notablemente conservados de un pez todavía más afín a los tetrápodos que *Panderichthys* o *Elpistostege*. La especie no se ha descrito ni denominado todavía, pero promete aportar datos fascinantes.

Respiración aérea

Gracias a estos hallazgos y a los análisis recientes, contamos con los restos de nueve géneros que registran unos 20 millones de años de la evolución temprana de los tetrápodos. Asimismo, arrojan luz sobre la adaptación del resto del cuerpo de los vertebrados a la vida en tierra firme. Revelan un aspecto evolutivo de sumo interés: lo mismo que en el desarrollo de las extremidades, muchas de las innovaciones decisivas surgieron mientras eran, en la mayoría de sus hábitos, acuáticos. Además, parece que los primeros cambios no estuvieron relacionados con la locomoción, sino con una dependencia creciente de la respiración aérea.

Este cambio de ventilación pudo haber iniciado las modificaciones



2. EN LA ORILLA: *Ichthyostega* corresponde al primer tetrápodo conocido que muestra adaptaciones para una locomoción no natatoria, aunque parece que su movimiento se asemejaba más al de una foca que al de un vertebrado terrestre. Poseía también rasgos acuáticos: una cola poderosa, patas posteriores en forma de aletas y un oído de uso subacuático, entre otros. No se sabe de qué manera *Ichthyostega* dividía su tiempo entre el medio terrestre y el acuático. Quizás excavaba nidos para sus huevos en tierra y cazaba y comía en el agua.

morfológicas graduales de la cintura escapular y las aletas pectorales. Se ha abordado la función de las protopatas en las formas de transición como *Acanthostega*, cuestionando que fuera la de locomoción. Las pruebas abonan la hipótesis siguiente: las aletas dirigidas hacia atrás ganaron en resistencia conforme se fueron convirtiendo en patas extendidas hacia los lados, con superficies extensas para la inserción muscular.

Aunque habrían de pasar millones de años antes de que las patas anteriores se desarrollaran hasta poder soportar el cuerpo sobre tierra, en el transcurso de dicho intervalo el animal podría haberlas utilizado para levantar la cabeza y respirar fuera del agua. Los dedos habrían participado en este movimiento facilitando la distribución del peso sobre las extremidades.

En 2004, el equipo de Shubin anunció un descubrimiento que reforzaba esta idea: el húmero de un tetrápodo de 365 millones de años. El hueso, hallado en Red Hill, un yacimiento fosilífero de Pennsylvania, parece haber estado unido al resto del cuerpo mediante una articulación de tipo charnela, en oposición a la articulación tipo rótula que los humanos y otros vertebrados terrestres poseemos. Esta disposición no habría permitido un paso de marcha, pero sí habría facilitado el tipo de elevación de la parte anterior del cuerpo que un tetrápodo hubiera necesitado para respirar fuera del agua. Podría haber ayudado también a mantener la postura del animal en el agua mientras se hallaba al acecho de presas.

La respiración aérea requería varios cambios en el cráneo y la mandíbula. En el cráneo, el hocico se alargó; los huesos que lo forman se redujeron en número y se suturaron entre sí, reforzando el hocico para que el animal lo levantara fuera del agua, en un medio que no ofrecía sostén. Se integraron los huesos de la parte posterior de la cabeza con una firmeza superior a la mostrada por los del resto del cráneo; proporcionaron así anclajes firmes para los músculos de la columna vertebral que elevan la cabeza en relación con el cuerpo.

La fusión de los huesos que constituyen la mandíbula inferior fortaleció esta estructura, facilitando la ventilación de "bomba bucal" que se supone en los tetrápodos. En este modo de respiración, empleado por los anfibios y los peces pulmonados modernos, la cavidad bucal se expande y se contrae como un fuelle para tragar aire e impulsarlo hacia los pulmones. Quizás el bombeo bucal exigió más potencia mandibular bajo la influencia de la gravedad que en el agua, donde los cuerpos son más o menos ingrávidos.

¿Se produjo el fortalecimiento de las mandíbulas, en cambio, para adaptarse a una alimentación terrestre? Es posible. Los primeros tetrápodos eran carnívoros; por tanto, resulta improbable que, de adultos, basaran su dieta en alimentos obtenidos en tierra firme, pues allí sólo habrían encontrado insectos y otros pequeños artrópodos. Pero los individuos juveniles necesitaban ese tipo de presas; quizá fueron ellos los que se aventuraron inicialmente en tierra firme para atraparlas.

Mientras tanto, desapareció, más atrás en el esqueleto, una serie de huesos que conecta la cabeza con la cintura escapular en los peces. De ahí que los tetrápodos posean, a diferencia de los peces, un cuello muscular que conecta la cabeza al resto del esqueleto y permite el movimiento de la cabeza independiente del cuerpo. También el sistema branquial experimentó una renovación importante: perdió algunos huesos, pero aumentó el tamaño del espiráculo (una abertura en la parte superior de la cabeza que conducía a un saco lleno de aire en la región gular), de forma que todo el aparato respiratorio estuviera mejor adaptado a la respiración aérea.

¿Por qué, después de millones de años de respirar con éxito bajo el agua, algunos peces empezaron a dirigirse a la atmósfera para obtener oxígeno? La forma del cráneo ha revelado algunas pistas sobre esta cuestión. Todos los tetrápodos primitivos y los cuasitetrápodos descubiertos hasta la fecha presentan un cráneo bastante plano (visto de frente). Esta



observación, combinada con datos paleoambientales recogidos de los yacimientos fosilíferos, nos remite a unos especialistas de aguas someras que acudían a los remansos para cazar peces de menor tamaño, aparearse y poner huevos. Durante el Devónico, las plantas vasculares prosperaban, quizá de manera no fortuita, transformando los dominios terrestres y los acuáticos. Por primera vez, las plantas caducifolias perdían las hojas según las estaciones cambiantes y creaban entornos atractivos para las presas pequeñas pero incómodos para la natación de los peces grandes. Además, debido a que el agua cálida contiene menos oxígeno que la fría, tales áreas habrían sido pobres en oxígeno. Así pues, los cambios en el esqueleto antes descritos pudieron haber facilitado a los tetrápodos primitivos el acceso a aguas a las que los tiburones y otros peces de gran tamaño no podían llegar, lo que les habría conferido una clara ventaja competitiva. Fue sólo casualidad que estos mismos rasgos resultaran útiles, andando el tiempo, para conquistar la tierra firme.

Estas innovaciones relacionadas con la respiración encaminaron a los tetrápodos por la senda que los convertiría en animales terrestres. Con todo, la vida en tierra requería ulteriores modificaciones del esqueleto; por ejemplo, una remodelación de la región auditiva. Se ignoran muchos detalles de esa transformación necesaria. Pero sí está claro que, incluso en los peces tetrapodiformes portadores todavía de aletas (*Panderichthys*, por ejemplo), se había acortado la parte del cráneo situada detrás de los ojos, tras previa reducción de la cápsula que alberga el oído interno. Si, como sugieren los datos paleoambientales, *Panderichthys* medraba en llanuras mareales someras o estuarios, la reducción del oído interno puede reflejar la creciente influencia de la gravedad en el sistema vestibular, que coordina el equilibrio y la orientación. Al propio tiempo, un aumento en el tamaño de la cámara de aire en la garganta pudo haber mejorado la audición. En algunos peces modernos, este saco aéreo "capta" ondas sonoras, evitando el paso directo de las mismas a través del cuerpo del animal. Desde allí se transmiten, mediante los huesos circundantes, al oído interno. Una cámara aérea engrosada, evidente en *Panderichthys*, habría interceptado mayor cantidad de ondas sonoras, aumentando así la capacidad auditiva.

Las modificaciones de la región auditiva guardaban también una estrecha relación con las del sistema branquial. En este sentido, la hiomandíbula (hueso que en los peces dirige los movimientos asociados a la alimentación y a la respiración) se redujo en tamaño y quedó alojada en un agujero de la cavidad craneana, donde se convirtió en el estribo. En los tetrápodos modernos, el estribo magnifica las ondas sonoras y las transmite del tímpano al oído interno a través del espacio aéreo de la garganta. (En los mamíferos, que poseen un sistema auditivo único, el estribo corresponde a uno de los tres huesecillos que constituyen el oído medio.) La primera etapa de la conversión debió transcurrir con prontitud, pues ya había culminado por la época de *Acanthostega*. Muy posiblemente, se dio a la par que el cambio de aletas a patas con dedos. El estribo, en cambio, tardaría millones de años en adquirir su función conocida en el oído timpánico adaptado a la vida terrestre. Entretanto, en estos tetrápodos todavía acuáticos operaría como un componente estructural del cráneo.

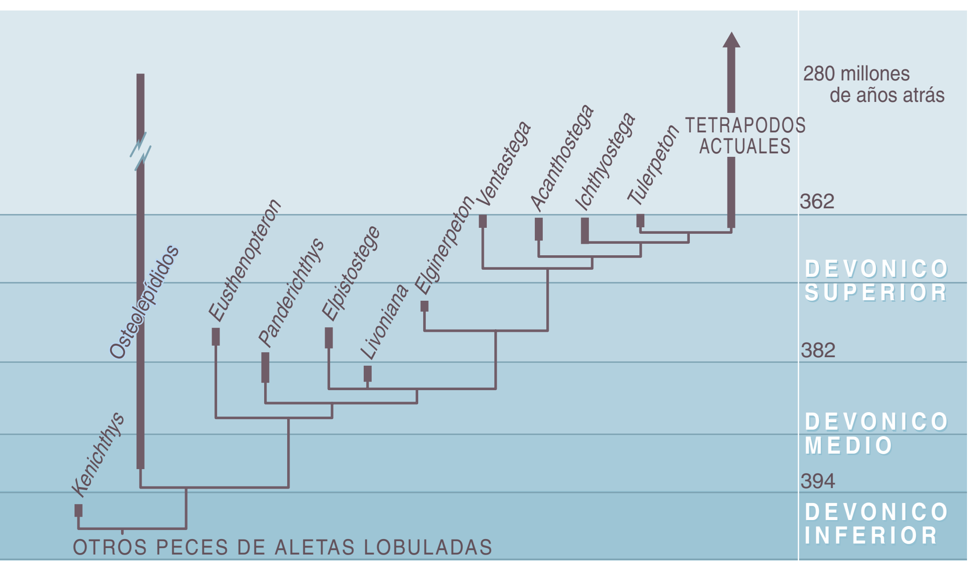
Tomadas en su conjunto, estas transformaciones del esqueleto nos han obligado a modificar la doctrina recibida sobre los tetrápodos primitivos. Las que antaño se consideraron obras en curso evolutivo (patas u oídos incompletamente desarrollados, por ejemplo), constituyeron adaptaciones por derecho propio. No siempre exitosas, pero adaptaciones. En cada estadio de esta transición hubo innovadores, algunos de ellos especializadísimos, que conquistaron nuevos nichos.

Biodiversidad

En conjunto, los tetrápodos y cuasitetrápodos dotados de patas desenterrados hasta la fecha eran animales de tamaño considerable, de alrededor de un metro de largo. Depredaban sobre invertebrados y peces, sin seleccionar demasiado sus presas. Aunque están comenzando a aparecer excepciones a esta regla generalista. Se sabe que *Livoniana*, descubierto en 2000 en un museo lituano por Per Erik Ahlberg, de la Universidad de Uppsala, porta, a tenor de los fragmentos de la mandíbula inferior conservados, siete hileras de dientes, no las dos habituales en cada lado de la mandíbula. Lo que *Livoniana* podía consumir con esta dentición de aspecto de mazorca no lo sabemos; pero lo más probable es que su dieta fuera distinta de la de sus parientes.

Contradicen la tesis recibida los nuevos trabajos sobre *Ichthyostega*, el primer tetrápodo conocido del

LUCY READING-IKKANDA



3. RELACIONES DE LOS TETRAPODOS. Los tetrápodos surgieron a partir de peces de aletas lobuladas como *Eusthenopteron* hace entre 380 y 375 millones de años, a finales del Devónico Medio.



Devónico: tampoco él se ceñía a la norma. La región auditiva y las partes asociadas de la bóveda craneana vienen desconcertando desde hace tiempo a los expertos porque presentan una construcción distinta de cualquier otro tetrápodo o pez conocido. Pero la identificación de nuevos fósiles, la obtención de nuevas preparaciones de material recolectado previamente y, sobre todo, el análisis mediante tomografía axial computerizada de especímenes clave, empiezan a arrojar luz sobre esta cuestión.

Según parece, *Ichthyostega* poseía un oído especializado, adaptado para operar bajo el agua. Carente de tímpano, lo que le aleja de la mayoría de los animales terrestres modernos, poseía, a cada lado de la parte posterior de la cabeza, una cámara con los tabiques superior y laterales reforzados, que probablemente estaba llena de aire. Sobre el suelo membranoso de la cámara se extendía un estribo muy delicado, en forma de cuchara, que debía vibrar en respuesta al sonido que incidía directamente en el aire de la cámara, para transmitir luego dichas vibraciones al oído interno a través de un agujero en la pared de la caja craneana. Ello implicaría que *Ichthyostega* pasaba gran parte de su tiempo en el agua. Asimismo, su aleta caudal y las patas posteriores tipo aletas sugieren un estilo de vida acuático.

Sin embargo, otras partes del esqueleto de *Ichthyostega* revelan una capacidad de desplazarse en tierra. Poseía escápulas y brazos increíblemente potentes. Las costillas de la región torácica, muy anchas, se superponían, formando un corsé que habría evitado que la cavidad torácica y los pulmones se aplastaran cuando se hallaba en tierra. Ello no obstante, *Ichthyostega* se encontraría todavía lejos de moverse con la soltura de un vertebrado terrestre. De entrada, su caja torácica habría restringido la ondulación lateral del tronco que típicamente tiene lugar en el movimiento de los tetrápodos. Además, en contraste con los peces, *Acanthostega* u otros tetrápodos primitivos, *Ichthyostega* poseía espinas sobre las vértebras que cambiaban de dirección a lo largo de la columna vertebral, lo que indicaría que los músculos que soportaban se hallaban especializados para la realización de tareas distintas y que se movía de una manera singular. Esta disposición multidireccional de las espinas vertebrales muestra un paralelismo con la de los mamíferos actuales, pero no se conocía en los tetrápodos del Devónico hasta que se estudió *Ichthyostega*. Las pruebas recabadas sugieren que, en vez de doblarse en el plano horizontal, como hace el cuerpo de un pez, el de *Ichthyostega* se doblaba principalmente en un plano vertical. Las patas traseras en forma de canalete no parecen haber contribuido demasiado al empuje hacia delante durante la locomoción: eran las robustas patas anteriores y sus escápulas poderosas las que lo proporcionaban. Así, en tierra *Ichthyostega* debía de moverse a la manera de una foca: primero levantando el dorso, después adelantando simultáneamente ambas extremidades anteriores y, por fin, tirando hacia delante del resto del cuerpo.

A tenor de los resultados de nuestro estudio, *Ichthyostega* constituye el vertebrado conocido más antiguo que

ANDREW RECHER (*tetrápodos*); LUCY READING-IKKANDA (*mapa*)

muestra adaptaciones a la locomoción no natatoria. No podemos saber con certeza qué hacía en la orilla. Quizás allí comía peces varados, pero se reproducía en el agua, donde habría usado su oído especializado para advertir la presencia de parejas potenciales. (Ello implicaría que, amén de percibir sonidos, *Ichthyostega* también los emitía.) O quizás *Ichthyostega* comía y detectaba (auditivamente) las presas en el agua, mientras que en tierra se servía de las patas anteriores para enterrar sus huevos. Con todo, su particular diseño corporal estaba sentenciado, pues no se ha hallado ningún fósil perteneciente a su estirpe posterior a hace 360 millones de años.

No cabe duda de que muchos otros diseños fracasaron en el curso de la evolución de los primeros tetrápodos. Aunque la confirmación de estas hipótesis requiere estudios adicionales, los últimos datos demuestran que entre los tetrápodos del Devónico existía mayor diversidad de lo que se creía.

Biogeografía y paleoecología

Los fósiles descubiertos en el curso de los últimos veinte años han arrojado luz sobre la historia de cambios acontecidos en el esqueleto de los tetrápodos. Así como sobre el lugar y el tiempo de su evolución. Aparecieron hace de 380 a 375 millones de años, a finales del Devónico Medio, una fecha bastante más ajustada que la que se había postulado con anterioridad. Los primeros representantes de este grupo eran cosmopolitas.

Los tetrápodos del Devónico estaban dispersos por todo el globo: desde lo que hoy es China y Australia, donde han aparecido *Sinostega* y *Metaxygnathus*, respectivamente, hasta los EE.UU. orientales, donde se hallaron *Hynerpeton* y el húmero de Red Hill. Si se sitúan los yacimientos en un mapa paleogeográfico de aquella época, se advierte que habitaban en todos los trópicos y subtrópicos de un supercontinente que comprendía Laurasia en el norte y Gondwana en el sur. La distribución casi ubicua en los climas más cálidos constituye una prueba del éxito evolutivo de estos animales.

Los tetrápodos del Devónico medraban, pues, en entornos muy dispares. Los yacimientos de Groenlandia oriental, los primeros en proporcionar fósiles de ellos, indican que la región fue antaño una extensa cuenca fluvial dominada por inundaciones periódicas que alternaban con estaciones más secas. El río, de agua dulce, ofrecía la base para los ambientes en los que se supone que evolucionaron los tetrápodos. Pero esa hipótesis quedó en entredicho tras el descubrimiento de *Ventastega*, *Tulerpeton* y otros animales en yacimientos de salinidad variable. Una zona fosilífera de especial fecundidad, Red Hill, en Pennsylvania, constituyó el lugar idóneo para el asentamiento de tetrápodos, pues allí medraban numerosas especies de peces, invertebrados y plantas. Lo mismo que los yacimientos de Groenlandia oriental, representa una cuenca fluvial. En cuanto al clima de la región, los estudios paleoambientales sugieren que no era monzónico (como se asocia a los hallazgos de Groenlandia), sino templado. Ello significa que la distribución de los tetrápodos primitivos quizás era más amplia de lo que se pensaba.

Investigaciones futuras

Queda mucho por aprender sobre los cambios anatómicos que acompañaron al progreso de los tetrápodos. Disponemos ya de un modelo razonable que permite explicar la evolución de la cintura escapular y las patas anteriores, pero carecemos de otro similar para el origen del robusto complejo de las patas posteriores (rasgo fundamental de los tetrápodos), pues los fósiles descubiertos hasta la fecha no ofrecen pista alguna. Sólo algunos ejemplares de *Ichthyostega* y *Acanthostega* conservan esta parte de la anatomía; en ambos, las patas posteriores están demasiado bien constituidas para revelar las fases precedentes. No existe ningún modelo que explique todos los estadios de la transición y nos falta un cuadro más ajustado del orden en que procedieron los cambios en el esqueleto (por ejemplo, cuándo evolucionó la extremidad posterior en relación a la anterior y al oído).

El descubrimiento y descripción de nuevos fósiles, con el recurso a la biología evolutiva del desarrollo, permitirán resolver algunos interrogantes. Realizados con ese propósito, hay ya estudios de los mecanismos de control genético que rigen la formación de la región branquial en los peces y el área del cuello en los mamíferos y las aves que están empezando a proporcionar pistas plausibles sobre procesos comunes de tetrápodos y peces y sobre procesos exclusivos de los tetrápodos.

Sabemos, en particular, que los tetrápodos han perdido todos los huesos que protegen las branquias en los peces; y, además, que los genes que gobiernan su formación los portan todavía los ratones, aunque con funciones distintas. Sabemos también que, en la región del cuello, se han desbaratado las vías bioquímicas encargadas del desarrollo de las patas. El crecimiento de extremidades adicionales se induce con facilidad (en el laboratorio) en el flanco de un tetrápodo, pero no en el cuello. Algo especial debió ocurrir cuando los tetrápodos desarrollaron un cuello, que impidió que allí brotaran patas.

Otras cuestiones quizás entrañen mayor complejidad. ¿Cuál de los numerosos ambientes en los que se han descubierto tetrápodos fósiles constituyó la cuna de los primerísimos miembros de este grupo? Las pruebas indican sólo que no debutaron en entornos estrictamente marinos. ¿Qué presiones evolutivas operaron en cada fase de la transición? Al carecer de un registro fósil completo, quizá no logremos ensamblar todas las piezas del rompecabezas de la evolución de los tetrápodos. Pero sí podemos aspirar a colmar algunas de las lagunas que quedan en la historia de la conquista de la tierra firme.

La autora

Jennifer A. Clack enseña paleontología de vertebrados en la Universidad de Cambridge. Lleva 25 años estudiando los orígenes de los tetrápodos.

Bibliografía complementaria

GAINING GROUND: THE ORIGIN AND EVOLUTION OF TETRAPODS. Jennifer A. Clack. Indiana University Press, 2002.

THE AXIAL SKELETON OF THE DEVONIAN TETRAPOD *ICHTHYOSTEGA*. Erik Ahlberg, Jennifer A. Clack y Henning Blom en *Nature*, vol. 437, págs. 137-140; 1 de septiembre 2005.

THE EMERGENCE OF EARLY TETRAPODS. Jennifer A. Clack en *Palaeogeography, Palaeoclimatology, Palaeoecology* (en prensa).

Medición de radiaciones en seres vivos

La dosimetría es una parte esencial de cualquier investigación sobre los posibles efectos biológicos de los campos electromagnéticos

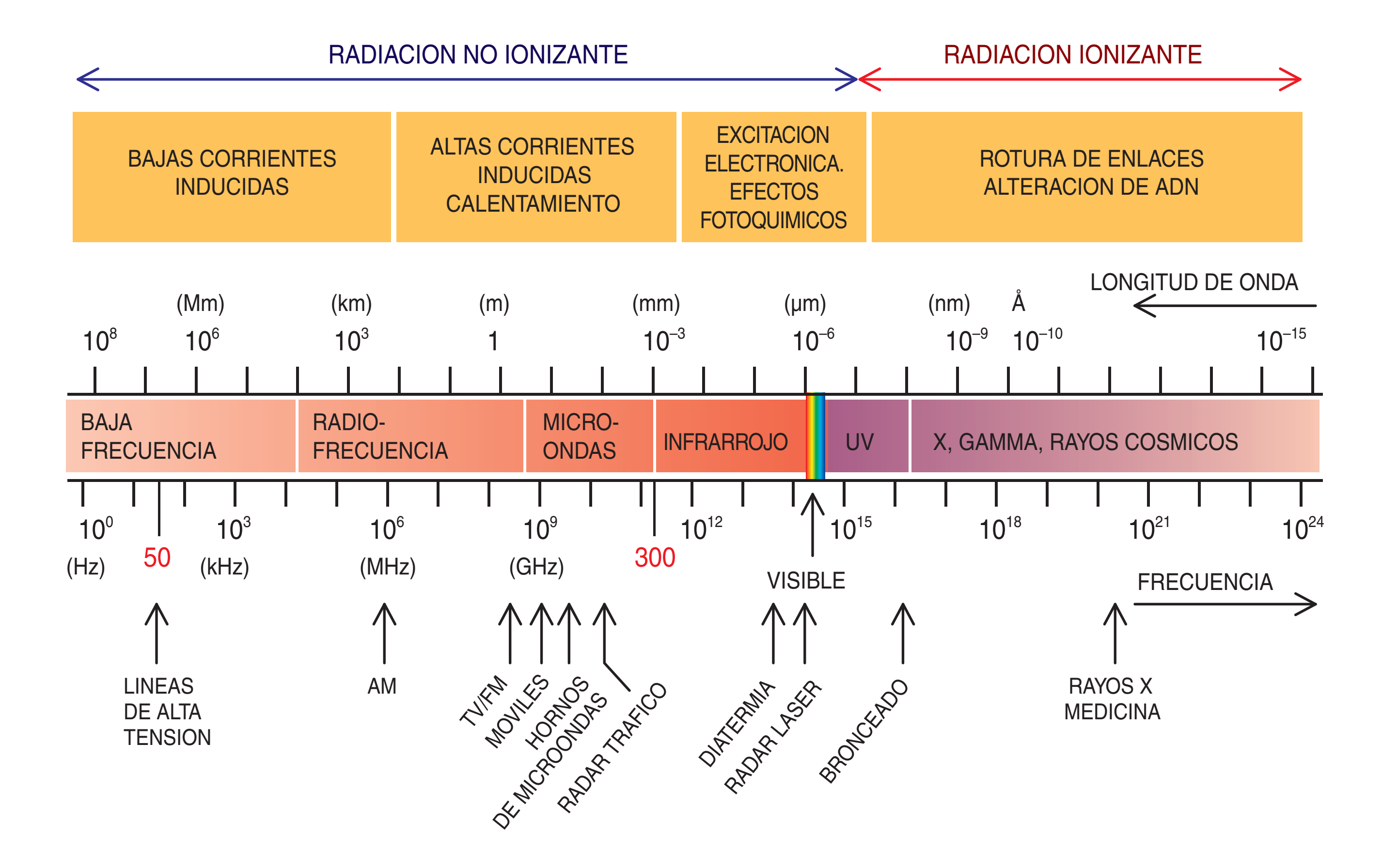
José Luis Sebastián Franco, Sagrario Muñoz San-Martín,
Miguel Sancho Ruiz y José Miguel Miranda Pantoja

El desarrollo, tan espectacular, de la telefonía móvil ha inundado la sociedad en que vivimos de energía electromagnética. Aunque la influencia de la electricidad y de los campos electromagnéticos en los sistemas biológicos ya fue objeto de atención en el siglo XVIII por Luigi Galvani y en el siglo XIX por Jacques-Arsène d'Arsonval, es a partir de la segunda guerra mundial cuando realmente se inician los estudios rigurosos de la interacción entre los seres vivos y la energía electromagnética. Hoy constituyen una de las líneas de investigación científica más activas en todo el mundo. El interés por estos estudios se despertó con la alarma que, en el decenio de 1960, surgió entre los trabajadores especializados en sistemas de radar y comunicaciones. Luego, el principal impulso ha venido de la inquietud de la población por el daño potencial que los campos de radiofrecuencia (RF) a que nos exponen los terminales de telefonía móvil y las antenas de las estaciones base en los edificios ejerzan sobre la salud.

Cuando se habla de los posibles efectos biológicos causados por la exposición a una radiación electromagnética, se suelen distinguir tres clases: *térmicos*, *atérmicos* y *no térmicos*. Los tres términos son relativos y por consiguiente cuesta delimitar la zona en que se solapan. Los *efectos térmicos* ocurren cuando en el organismo se deposita la suficiente energía de radiofrecuencia como para que experimente un incremento de temperatura medible, por ejemplo de 0,1 ºC. Los *efectos atérmicos* se producen cuando se deposita la suficiente energía para que haya un incremento de temperatura en el objeto biológico, pero sin que, gracias a sus procesos termorreguladores, llegue a ser considerable. Finalmente, los *efectos no térmicos* son los que se dan cuando la energía depositada en el objeto biológico no tiende a producir aumento de temperatura, incluso prescindiendo de posibles mecanismos termorreguladores. Por un principio de precaución, se han establecido recomendaciones que imponen fuertes restricciones a los efectos térmicos de la exposición a una radiación de RF. Los criterios que se han seguido para establecer estas normas no han sido los mismos en todos los países. Así, en las naciones orientales los criterios se fundan en la respuesta del sistema nervioso central y en el comportamiento de los animales después de que se los exponga a campos de radiofrecuencia. Sin embargo, en Europa y Estados Unidos se han basado casi exclusivamente en los efectos térmicos que se producen en los seres humanos expuestos a radiofrecuencias. La diversidad de criterios hace necesario que se sigan investigando métodos más perfectos que relacionen los efectos observados en animales con los encontrados en personas. De igual modo, estas normas de seguridad se deben revisar y modificar a medida que se disponga de mayores datos sobre los efectos

Distintas bandas de frecuencia y fuentes en el intervalo 0-300 GHz

BANDA	NOMENCLATURA	INTERVALO	FUENTES
Extremadamente alta	EHF	30-300 GHz	Radar, exploración espacial
Super alta	SHF	3-30 GHz	Radar, comunicación satélites
Ultra alta	UHF	300-3000 MHz	Radar. TV, móviles, industrial
Muy alta	VHF	30-300 MHz	TV, FM
Alta	HF	3-30 MHz	SW radio, diatermia onda corta
Media	MF	300-3000 KHz	AM, radio marítima
Baja	LF	30-300 KHz	Navegación, radio faro
Muy baja	VLF	3-30 KHz	Navegación, sonar
Ultra baja	ULF	300-3000 Hz	Rango de audio
Super baja	SLF	30-300 Hz	Comunicación submarina, red



1. ESPECTRO ELECTROMAGNETICO. El intervalo de frecuencias 0-300 GHz, que incluye la mayor parte de los campos electromagnéticos ambientales, está dentro del rango correspondiente a las radiaciones no ionizantes.

JOSE LUIS SEBASTIAN FRANCO, SAGRARIO MUÑOZ SAN-MARTIN, MIGUEL SANCHO RUIZ Y JOSE MIGUEL MIRANDA PANTOJA

de los campos de radiofrecuencia en el cuerpo humano.

Frecuencias

Los estándares modernos de seguridad en las exposiciones a campos electromagnéticos abarcan el intervalo de frecuencias desde 0 hertz (Hz) hasta 300 gigahertz (GHz). Además de los campos externos, hay que considerar también los endógenos. En la membrana celular existe una diferencia de potencial del orden de 100 milivolt, lo que da lugar a un campo eléctrico muy intenso. Un impulso nervioso puede transportar densidades de corriente próximas a un ampère por metro cuadrado; en consecuencia, el cerebro y el corazón establecen campos magnéticos del orden de 0,5 nanotesla y 5 picotesla, respectivamente, con frecuencias comprendidas entre 0,1 y 60 Hz.

El intervalo de frecuencias de 0 a 300 GHz cae dentro del rango correspondiente a las radiaciones no ionizantes, es decir, la energía de un fotón dentro de este intervalo es *insuficiente para alterar* la estructura interna de las moléculas o *romper enlaces* intermoleculares. A diferencia de una radiación ionizante, una radiación no ionizante produce efectos biológicos muy distintos en diferentes especies animales o cultivos según cuál sea la frecuencia. El intervalo de frecuencias 0-300 GHz es también muy peculiar: en la parte baja de este intervalo las longitudes de onda son mucho mayores que las dimensiones de la estructura atómica o molecular y en la parte alta suelen ser menores que las dimensiones de los órganos y entidades biológicos. Cada una de estas circunstancias permite simplificar el cálculo de los campos mediante aproximaciones.

Campos internos

La interacción de una estructura biológica con una radiación electromagnética es complejísima. Ocurren procesos no lineales e irreversibles en el objeto biológico. Por otra parte,

En el texto se define la permitividad como la constante de proporcionalidad entre el campo aplicado externo y el campo en el interior de un material. Con esta definición, lo natural es que la permitividad del vacío valga uno, tal y como ocurre en el sistema cgs de unidades; sin embargo, en el Sistema Internacional se prefirió que la carga eléctrica fuese una magnitud irreducible a otras, lo que lleva a que la permitividad adquiera unidades, en vez de ser un número puro; entonces, la permitividad del vacío ya no vale uno. Para convertir los valores de las permitividades de sustancias que se darán aquí en los correspondientes valores del Sistema Internacional, hay que multiplicarlos por el de la permitividad del vacío en éste, $8{,}8542 \times 10^{-12}$ C^2/Nm^2. Las demás unidades que se utilizan en el artículo son las del Sistema Internacional.

La Redacción

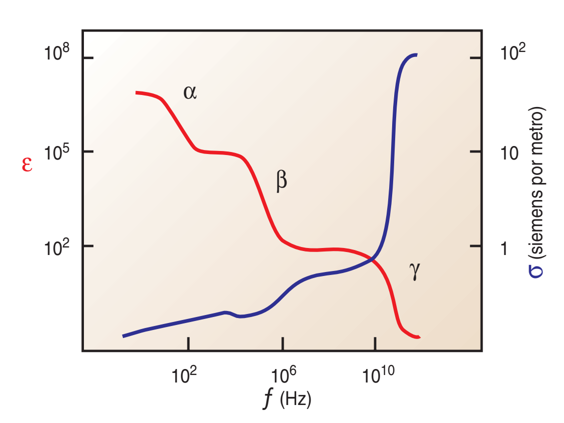
las propiedades eléctricas —la permitividad y la conductividad— de sus componentes, que determinan finalmente la distribución de los campos eléctrico y magnético en su interior, dependen de la frecuencia de la radiación.

Al aplicar un campo eléctrico de origen externo a un material dieléctrico, es decir, mal conductor, éste se polariza: las cargas eléctricas de sus moléculas se separan, positivas y negativas, y tienden a alinearse con ese campo. Esos dipolos así creados, con sus correspondientes campos asociados, contrarrestan en parte el campo aplicado. En algunas circunstancias, el campo resultante dentro del material es proporcional al campo aplicado externo. La permitividad del material es la correspondiente razón entre éste y aquél. La conductividad refleja la facilidad con que los iones se desplazan por el medio impulsados por el campo eléctrico.

Los fenómenos de polarización y conducción no son instantáneos; los efectos del campo dependerán, pues, de su frecuencia, el ritmo a que varía su intensidad y se invierte su sentido. Si esos cambios son demasiado rápidos, los dipolos no podrán seguirlos y su efecto atenuador del campo aplicado en el interior del tejido disminuirá, así que también disminuirá la permitividad. El comportamiento de la conductividad en función de la frecuencia del campo aplicado es el contrario: tiende a aumentar.

El rango de frecuencias que hay que considerar es muy amplio, de las decenas de hertz de una línea de alta tensión a los miles de millones de hertz de las microondas de un horno o de la telefonía móvil. A lo largo de todos esos órdenes de magnitud de la frecuencia, la permitividad de los tejidos orgánicos disminuye en casi otros tantos órdenes de magnitud. La conductividad, en cambio, aumenta sólo unos dos órdenes de magnitud.

La tendencia general de la permitividad es a disminuir a medida que aumenta la frecuencia. Presenta una forma característica para algunos intervalos de la frecuencia, denominada "zona de relajación", que refleja la reacción de determinados componentes de la estructura de las células a la frecuencia. Se aprecian tres de esos intervalos de frecuencia donde la permitividad disminuye de forma apreciable con la frecuencia. Se los identifica con las letras α, β y γ. La zona α se debe fundamentalmente a la contribución de los iones próximos a la célula, atraídos por las cargas de distinto signo que ésta tiene en su superficie, y a la polarización de estructuras celulares conectadas a la membrana plasmática. La zona β se debe a la transferencia de carga en la membrana celular y los orgánulos celulares (las mitocondrias, el núcleo de la célula). Por último, la zona γ corresponde a las rotaciones de los dipolos del agua libre y de la ligada en los tejidos, que a otras longitudes de onda no se producen, estorbadas por los enlaces que ligan unas moléculas con otras; se extiende desde unos pocos megahertz hasta unos 20 GHz. La región de las microondas, pues, cae casi entera más allá de la zona gamma, que sin embargo incluye todavía las microondas utilizadas en los hornos y en la telefonía móvil.



2. LA PERMITIVIDAD (ε) Y LA CONDUCTIVIDAD ELECTRICA (σ) de un tejido disminuye y aumenta, respectivamente, a medida que la frecuencia de un campo eléctrico oscilante aplicado aumenta. El cambio abarca muchos órdenes de magnitud. Para determinados intervalos de frecuencias —letras α, β y γ— se aprecian en la curva de la permitividad unos suaves escalones denominados "zonas de relajación".

Dosimetría

El punto de partida para estudiar la interacción entre la radiación y el objeto biológico debe ser la determinación de la energía absorbida por un cuerpo expuesto a los campos eléctrico y magnético que componen una señal de RF: la *dosimetría*. Esta energía absorbida está directamente relacionada con los campos "internos": los campos eléctrico y magnético dentro del objeto. Son muy distintos de los incidentes; dependen del tamaño y la forma del objeto, de sus propiedades eléctricas, de su orientación respecto a los campos incidentes y de la frecuencia. Sin embargo, puesto que la medida directa de los campos incidentes es mucho más fácil de realizar que la de los campos internos, la dosimetría resulta esencial para relacionar los campos internos, que son los que producen el posible efecto, con los campos incidentes.

La dosimetría es de especial importancia cuando se realizan experimentos con animales pequeños, ya que el tamaño es un factor fundamental en la cantidad de energía absorbida. Como ejemplo, la energía media absorbida por unidad de masa por una rata de tamaño medio a la frecuencia de uso industrial de 2,45 GHz llega a ser diez veces superior a la absorbida por una persona de tamaño medio expuesta a los mismos campos incidentes. De igual forma, debido a que los campos internos pueden ser distintos, un efecto biológico observado en un animal expuesto a una radiación de RF quizá no se manifieste en otro animal de tamaño distinto y expuesto a la misma radiación incidente.

Los datos dosimétricos se presentan en la forma de una "tasa de absorción específica" (SAR, acrónimo de *specific absortion rate*), que se expresa en watt por kilogramo. El término SAR fue propuesto originalmente en 1981 por el Consejo Nacional para la Medida y la Protección de la Radiación (NRCP); el Instituto Nacional Americano de Estándares (ANSI) sería el primero en adoptarlo, y luego se extendió a toda la comunidad científica. SAR es el ritmo al que un cuerpo absorbe energía cuando está expuesto a una radiación de RF. Es proporcional al cuadrado de la intensidad del campo en el interior del tejido; la constante de proporcionalidad es igual a la conductividad dividida por la densidad del tejido.

JOSE LUIS SEBASTIAN FRANCO, SAGRARIO MUÑOZ SAN-MARTIN, MIGUEL SANCHO RUIZ Y JOSE MIGUEL MIRANDA PANTOJA

Las propiedades eléctricas de los tejidos (permitividad mucho mayor que la del vacío y conductividad apreciable) determinan la velocidad de propagación de una onda electromagnética en su interior, siempre menor que en el espacio vacío. Además, dentro de un tejido el valor de la razón entre los campos magnético y eléctrico es mucho más baja que en el espacio vacío. Como resultado, parte de la energía incidente se refleja (al igual que ocurriría en la transmisión de señales al unir dos cables con características eléctricas distintas), se reduce la longitud de onda de la radiación en el tejido y aumenta el ritmo de atenuación del campo electromagnético a medida que penetra en el tejido. Es evidente entonces que la cantidad de energía absorbida y la forma en que se distribuye espacialmente dentro del tejido biológico van a ser peculiares de cada tejido y dependientes de sus propiedades eléctricas. Por consiguiente, la caracterización eléctrica de los diferentes tejidos a distintas frecuencias de radiación es un aspecto fundamental para la dosimetría.

Sin embargo, es muy difícil encontrar en la bibliografía especializada valores experimentales únicos para la permitividad y la conductividad de un tejido. Esta falta de reproducibilidad en los valores se debe a varias causas: las diferentes formas en las que se han obtenido las muestras, las distintas técnicas experimentales empleadas, la falta de control de las condiciones en que se ha realizado la medida y otras.

Las propiedades de los distintos tejidos se pueden describir en gran medida por su contenido de agua, que es el que condiciona los valores de la permitividad y conductividad. Así, el ojo, músculo, piel, hígado y riñón tienen un alto contenido de agua mientras que en la grasa y el hueso es muy bajo. Esto puede explicar el amplio margen de los valores de permitividad (entre 5 y 80) que se han medido a la frecuencia de 30 GHz. Los valores más altos corresponden a la sangre, hígado, músculo (alrededor de 50) y cerebro (alrededor de 40), y los más bajos (por orden decreciente) al hueso, la piel y la grasa (alrededor de 5). Del mismo modo, para la conductividad se han medido valores comprendidos entre 0,3 y 2,5 siemens por metro (el siemens por metro, S/m, es la unidad de conductividad del Sistema Internacional de unidades de medida; un siemens es 1/ohmio), también para 30 GHz.

Además de las propiedades eléctricas del tejido biológico, existe una serie de factores que también desarrollan un papel importante en el cálculo de la energía absorbida por el tejido, como son su geometría y tamaño, su orientación respecto del campo incidente, la polarización del campo (es decir, la dirección en que oscila el campo, que no debe confundirse con la polarización eléctrica del material, aunque la orientación de ésta depende de la de aquélla), la frecuencia, el tipo de fuente de la radiación, el entorno en que se produce la exposición y finalmente el tiempo de exposición a la radiación incidente, su intensidad y cómo esté modulada.

Pero centraremos nuestra atención en cómo se distribuyen los campos relacionados con esa absorción de energía dentro del objeto, tanto a escala macroscópica como microscópica.

Consideraremos primero el campo eléctrico. Como ya se ha mencionado anteriormente, gran parte del contenido de los tejidos humanos está constituido por agua (cerca del 85 %), con una concentración relativamente alta de iones. Al ser el agua un medio conductor, los iones se desplazan libremente hasta la superficie del cuerpo. Crean así una distribución de carga que anula el campo eléctrico en el interior; es decir, el cuerpo humano apantalla el campo eléctrico.

Para calcular la distribución a frecuencias bajas, por ejemplo a 50 Hz, la frecuencia de la red de distribución de energía eléctrica, consideraremos que la exposición se realiza en el espacio libre y que la conductividad del cuerpo humano tiene un valor medio de 0,5 siemens por metro. La intensidad del campo interno es del orden de las cienmillonésimas del externo; es decir, a estas frecuencias tan bajas, el campo eléctrico dentro del cuerpo humano tiene un valor prácticamente nulo.

El campo se va atenuando de manera exponencial a medida que penetra en el tejido. El ritmo lo determina un parámetro δ, llamado profundidad de penetración: a una distancia igual a δ, el campo se ha reducido en un factor 1/e (0,37). Este parámetro δ es inversamente proporcional a la raíz cuadrada de la frecuencia de la radiación y de la conductividad del tejido: para una conductividad de 0,5 S/m, vale 0,72 metros a la frecuencia de 1 MHz; pero a 1 GHz se reduce a 2,25 centímetros.

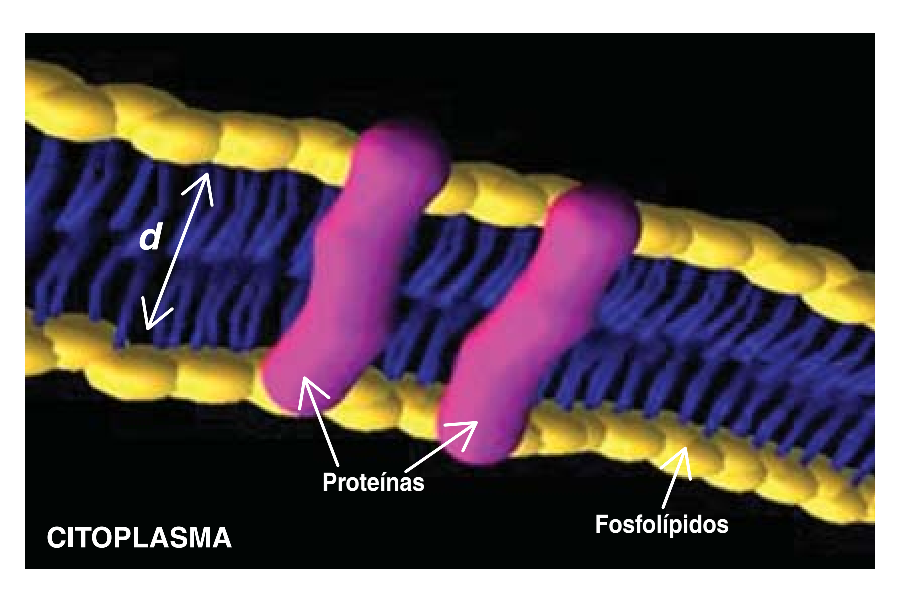
Los campos en las células

En estos últimos párrafos hemos centrado nuestra atención en la energía media absorbida por un objeto biológico, un tejido por ejemplo, sin considerar los detalles microscópicos de su estructura. ¿Qué ocurre en las células? ¿Cuáles son los posibles efectos de una radiación sobre sus distintos componentes? ¿Cómo se distribuyen los campos en su interior? Para contestar a estas preguntas es necesario realizar una simulación mediante técnicas numéricas y a escala microscópica de la interacción entre campo y una célula.

El principio básico de las técnicas numéricas que se aplican es el mallado o división en celdas del objeto expuesto. Se determina el campo eléctrico en el centro de cada celda por un método iterativo de cálculo. En estructuras complejas como la célula, donde se combinan

3. ATENUACION DEL CAMPO ELECTRICO a medida que profundiza en un tejido formado por piel, grasa y hueso con distintos espesores y valores de conductividad.

JOSE LUIS SEBASTIAN FRANCO, SAGRARIO MUÑOZ SAN-MARTIN, MIGUEL SANCHO RUIZ Y JOSE MIGUEL MIRANDA PANTOJA



4. LA MEMBRANA QUE ENCIERRA EL CITOPLASMA consta de fosfolípidos y proteínas. Sumergida la célula en un electrolito y expuesta a una radiación electromagnética, el campo eléctrico se amplifica en el interior de la membrana respecto al campo incidente.

distintas escalas (nanómetros para la membrana y micras para el citoplasma), debe utilizarse un mallado variable en las distintas partes de la estructura. Así, el mallado de la membrana ha de ser mucho más fino que el del citoplasma. La determinación de campos electromagnéticos mediante simulación por técnicas numéricas constituye actualmente una de las herramientas más potentes y útiles en la disciplina del bioelectromagnetismo.

Sin embargo, con la aplicación de algunos conceptos básicos de electromagnetismo cabe obtener de forma aproximada la distribución de campos dentro de una célula sin necesidad de recurrir a técnicas numéricas tan complejas. Consideremos un modelo sencillo, en el que la célula está inmersa en un electrolito con valores ordinarios de la permitividad, alrededor de 50, y de la conductividad eléctrica, medio siemens por metro. La célula está formada por el citoplasma, de ordinario con un valor típico de 10 micras de radio, rodeado por una membrana, constituida por fosfolípidos y proteínas y con un espesor "*d*" de entre 5 y 10 nanómetros. Para una frecuencia de 50 Hz, los valores de la permitividad y la conductividad de la membrana puede ser, respectivamente, 10 y 10^{-7} siemens por metro.

El campo que se induce a frecuencias bajas en el interior de un objeto que tiene una conductividad apreciable no depende de la polarización, sino, sobre todo, de la corriente que se genera en su interior y, por lo tanto, de su conductividad. En concreto, según las leyes de Maxwell, la razón entre la intensidad del campo dentro del electrolito y el campo externo aplicado es proporcional a la razón de la frecuencia de éste y la conductividad del electrolito, multiplicada por 3 al promediar para las tres direcciones del espacio. Con una frecuencia de 50 Hz y un electrolito de las características indicadas, la intensidad del campo en el electrolito será del orden de unas cienmillonésimas de la que habría tenido en el aire.

Con el mismo modelo, se obtiene, en cambio, que la intensidad del campo dentro de la membrana multiplica por 3000, más o menos, la intensidad del campo externo, debido al pequeño espesor y baja permitividad de la membrana. (Cuando esta permitividad es mucho menor que la razón entre la conductividad del electrolito y la frecuencia, el campo en la membrana depende sólo de la geometría de la célula: es inversamente proporcional al espesor de la membrana.) De igual modo, cuando se calcula el valor del campo dentro del citoplasma se obtiene un valor ligeramente superior al del electrolito. Por lo tanto, el valor del campo en la membrana es muy superior al campo en el electrolito y en el citoplasma: se ha producido un efecto de amplificación de campo en la membrana.

Esta amplificación de campo también se produce a altas frecuencias, aunque en menor grado. La membrana es, por consiguiente, el sitio preferente de interacción para la búsqueda de posibles efectos biológicos de los campos electromagnéticos. Existen además importantes efectos ligados a la forma geométrica y la proximidad de otras células que contribuyen también a esta amplificación de campo en la membrana. Para estudiarlos es necesario el uso de alguna de las técnicas numéricas anteriormente mencionadas.

A diferencia del campo eléctrico, la conductividad del objeto no apantalla el campo magnético. Así, por ejemplo, los posibles efectos biológicos debidos a las líneas de distribución de energía eléctrica a la frecuencia de 50 Hz (en Europa) se deberían fundamentalmente a los campos magnéticos generados por las corrientes que circulan por los cables. La intensidad de estos campos magnéticos decrece como el inverso de la distancia al cable. Sin embargo, a medida que la frecuencia aumenta, la variación del campo magnético con el tiempo genera un campo eléctrico, cuyos efectos hay que considerar.

Fantomas

Para determinar experimentalmente los valores locales de la tasa de absorción se pueden seguir dos procedimientos. El primero está basado en la medida de la intensidad del campo eléctrico en el punto de interés. Dicha tasa está especificada una vez que se conocen el valor del campo eléctrico, la conductividad eléctrica y la densidad del tejido. El segundo método consiste en la medición de variaciones de temperatura. Este segundo método se aplica fundamentalmente a tejidos y a cultivos celulares *in vitro*.

La tasa de absorción se puede determinar en cadáveres de animales, pero en general es más conveniente realizar la medida en un fantoma; resulta más reproducible y más exacta. Un fantoma es un material sintético que tiene, a las frecuencias de interés, unas propiedades eléctricas equivalentes a las del tejido biológico real en el que se quiere medir. Existen distintas fórmulas y procedimientos para preparar fantomas de la piel, el

JOSE LUIS SEBASTIAN FRANCO, SAGRARIO MUÑOZ SAN-MARTIN, MIGUEL SANCHO RUIZ Y JOSE MIGUEL MIRANDA PANTOJA

cerebro o el hueso a distintas frecuencias de radiofrecuencia. Se los puede moldear con la geometría del objeto biológico real. Para determinar la intensidad de campo eléctrico en un punto, se utiliza una sonda en miniatura de campo eléctrico o una sonda térmica implantada en el tejido o en el fantoma. La lectura de datos se realiza por medio de fibra óptica. Las propiedades térmicas de los fantomas determinan con qué rapidez la difusión térmica producirá cambios en los patrones de temperatura del material. Para tener una menor difusión térmica y que los patrones no se difuminen, en algunos fantomas se utiliza titanato de bario.

Las ventajas de utilizar sondas eléctricas implantables es la obtención directa del valor del campo eléctrico en un punto. Desplazando la sonda se obtiene información sobre la distribución de campo en una determinada región. Una sonda eléctrica está formada por tres hilos metálicos muy cortos, denominados antenas o dipolos, situados en las tres direcciones del espacio para obtener medidas isotrópicas (es decir, sin privilegiar ninguna dirección). Esto es importante debido a que la polarización del campo en el interior, en principio, es desconocida. La longitud l_{dip} de cada uno de los dipolos de la sonda desempeña una función fundamental, ya que la sensibilidad y la resolución espacial de la sonda son proporcionales a dicha longitud o a su inverso, respectivamente. Las sondas con l_{dip} que no supere los 5 mm pueden detectar valores de la tasa de absorción específica (SAR) de menos de 0,2 watt por kilogramo.

Sin embargo, con una fuente de microondas modulada en onda cuadrada a 1 kHz (una señal que alterna mil veces por segundo entre dos valores constantes) y un amplificador que actúa sobre esa misma frecuencia para determinar los valores máximo y mínimo del campo eléctrico en el cuerpo, se pueden medir SAR de sólo 0,2 mW/kg. Cuando la intensidad de campo eléctrico es muy débil, se deben utilizar otras técnicas para aumentar la sensibilidad, es decir, para distinguir mejor la señal en medio del ruido electromagnético presente en cualquier medición.

Cuando se hacen medidas a frecuencias bajas, las señales espurias que captan los cables de alta impedancia pueden degradar la resolución. Esto hace que las sondas eléctricas deban emplearse por encima de 150 MHz. El límite de frecuencia superior lo impone la razón de la longitud de onda de la radiación de RF en el fantoma y la longitud de dipolo de la sonda. Cuando no llega a 10, el campo eléctrico no tiene suficiente uniformidad en toda la longitud del dipolo y se degrada la precisión de la lectura. Como la longitud de onda en el tejido es siempre mucho menor que en el espacio vacío, el límite práctico de las sondas eléctricas implantables es de unos pocos GHz.

Sondas de temperatura

La determinación de distribuciones de SAR en fantomas, cadáveres y animales vivos ha experimentado un gran auge con el desarrollo de las sondas de temperatura inmunes a la radiofrecuencia. Los sistemas actuales de alta resolución, que utilizan detectores basados en termistores —resistencias variables con la temperatura— con hilos de alta impedancia, pueden determinar con exactitud cambios de temperatura de tan sólo 0,01 ºC en cuerpos biológicos expuestos a radiaciones intensas de RF. Cuando un tejido está expuesto a una radiación muy intensa durante un intervalo breve de tiempo, su temperatura crece linealmente. Como consecuencia, se pueden realizar medidas de tasas de absorción específica a partir de mediciones de temperatura. Sin embargo, la conductividad de un tejido aumenta con la temperatura, lo que provoca a su vez que se absorba más energía. Así, las áreas calientes absorben más energía de radiofrecuencia que las áreas frías. Con objeto de prevenir esta deriva térmica descontrolada, las variaciones máximas de temperatura se mantienen inferiores a 10 ºC. En estas condiciones la conductividad del tejido aumenta aproximadamente sólo un 2 % por grado centígrado. Aunque no ocurra la deriva térmica, los cambios en las propiedades dieléctricas del objeto biológico alteran la distribución de la energía absorbida. Esto hace que la exposición a una radiación de RF intensa deba ser de corta duración para minimizar la difusión de calor y que el incremento de temperatura no supere los 10 ºC.

La medida de la variación de la temperatura es el método que generalmente se emplea para determinar la tasa de absorción específica en tejidos y en cultivos celulares *in vitro*. Esto no quiere decir que el mecanismo de cualquier posible efecto que se pueda observar sea de origen térmico. La tasa de absorción específica es proporcional al calor específico del tejido y al incremento de temperatura, e inversamente proporcional al tiempo de exposición. La exactitud de esta relación depende de que las pérdidas térmicas por conducción del objeto irradiado no sean significativas. Por consiguiente, la irradiación del fantoma debe efectuarse con una señal de





CORTESIA DE SEMCAD

5. DOS MODELOS para estudiar los efectos de las radiaciones con frecuencias de radio en los seres humanos: un maniquí o fantoma antropomórfico (*izquierda*) correspondiente a un 90 % de las dimensiones de la cabeza de un adulto y fantoma numérico (*derecha*) de láminas de 1,1 mm que incluye humor vítreo, córnea, cerebro, cartílago, hueso, músculo, grasa y piel.

RF intensa y durante un tiempo corto, para así producir un incremento lineal de temperatura que se pueda medir. Si hay pérdidas térmicas apreciables, el valor de SAR obtenido será muy inferior al valor correcto.

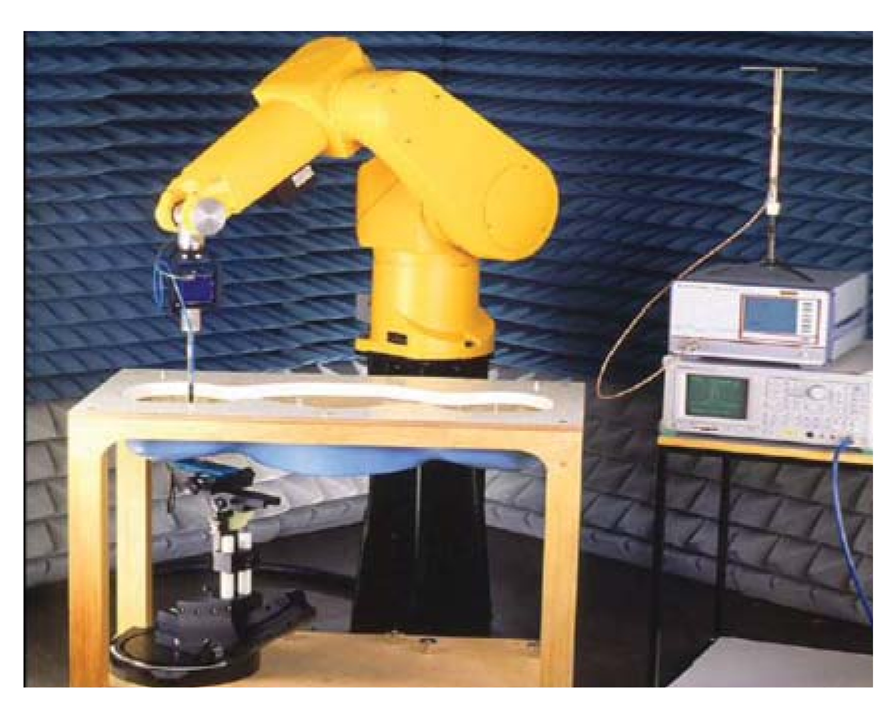
La exposición a RF de baja densidad de potencia durante intervalos muy largos de tiempo constituye un error muy común en la medida de la SAR, ya que la eliminación de energía térmica por medio de la difusión disminuye el aumento de temperatura inducido por la radiofrecuencia, con lo que se distorsiona de forma muy notable el valor real de la energía absorbida. Además, en los objetos biológicos vivos hay que tener en cuenta los efectos termorreguladores de la circulación sanguínea, que también influyen en el valor de SAR medido. Para una medida precisa de la tasa de absorción es fundamental determinar la tasa inicial de crecimiento de la temperatura.

Las sondas de temperatura tienen sin embargo sus limitaciones. Los sensores basados en termopares —termómetros que miden los voltajes generados por el calor en la soldadura de dos metales distintos— perturban la distribución original del campo de RF y por consiguiente no son adecuados para medidas precisas de SAR. Es necesario apantallar los termopares y termistores para mejorar la compatibilidad electromagnética con la radiación de radiofrecuencia.

Cuando se utilizan sistemas de medida que emplean cables con pérdidas y los campos son muy intensos, se pueden quemar los cables. Por ese motivo, estos sistemas funcionan mejor a frecuencias superiores a los 100 MHz. La estructura de una sonda de temperatura moderna consiste en una fibra óptica de diámetro inferior a 1 mm, con su extremo recubierto por una fina película de material sensible (fósforo) que actúa como sensor. Al ponerse en contacto con la superficie del cuerpo, el sensor reacciona muy rápidamente, debido a su pequeña masa térmica, alcanzando la temperatura de la superficie en un tiempo inferior a 100 milisegundos. La ventaja de utilizar fibra óptica para la transmisión de datos estriba en su inmunidad a las interferencias en el espectro de radiofrecuencia.

Existen numerosos trabajos experimentales en los que se utilizan sondas de temperatura para determinar el valor de la tasa de absorción específica local en modelos de la cabeza o del cuerpo humano expuestos a radiaciones muy intensas de radiofrecuencia. Se ha encontrado que para poder medir con precisión la tasa de absorción específica mediante una sonda térmica es necesario que existan puntos en el objeto biológico con valores de SAR de al menos 20 watt por kilogramo. Cuando con estos mismos modelos se utilizan fuentes de radiación de radiofrecuencia de baja potencia es muy difícil hacer medidas precisas con sondas térmicas. Por último, la sensibilidad de las sondas térmicas es inferior a 0,2 miliwatt por kilogramo. Ese valor es menor que el correspondiente a las sondas eléctricas.

En general, las medidas de valores locales de SAR son muy tediosas de realizar. Se requiere mucho tiempo para obtener una distribución espacial de la tasa de absorción específica local en un objeto y éste hay que enfriarlo entre exposiciones debido a que las propiedades eléctricas del objeto biológico dependen de su temperatura. Además, las sondas locales no garantizan la determinación de puntos calientes. Para ello, es necesario realizar un mapa detallado de valores de distribución de la tasa de absorción específica.



6. EL SISTEMA EXPERIMENTAL DASY4 de medida de la tasa de absorción específica local utiliza un fantoma y un robot para desplazar y situar una sonda en la posición deseada. La fuente de energía de radiofrecuencias es un terminal de telefonía móvil y los datos se registran en el medidor de campo situado sobre la mesa.

CORTESIA DE SEMCAD

Técnicas

Existen numerosas técnicas para la determinación de la distribución de la tasa de absorción específica en un plano dentro de un volumen tridimensional. Entre ellas se pueden destacar la lámina de cristal líquido sensible a la temperatura, la cámara termográfica por infrarrojos, las imágenes por resonancia magnética y la técnica de imágenes por luminiscencia. La lámina de cristal líquido proporciona una visualización *cualitativa* de la distribución de la temperatura. Esta lámina se introduce en un corte que previamente se ha realizado en el objeto y se deja durante la exposición a la radiación. A pesar de la sencillez de este método, su utilidad viene limitada por el carácter cualitativo de los resultados que proporciona.

Con objeto de obtener varios cientos de miles de puntos simultáneamente en cuestión de segundos se desarrolló la técnica termográfica. En ella se utilizan fantomas o animales que se diseccionan en los distintos planos de interés. Después de expo-

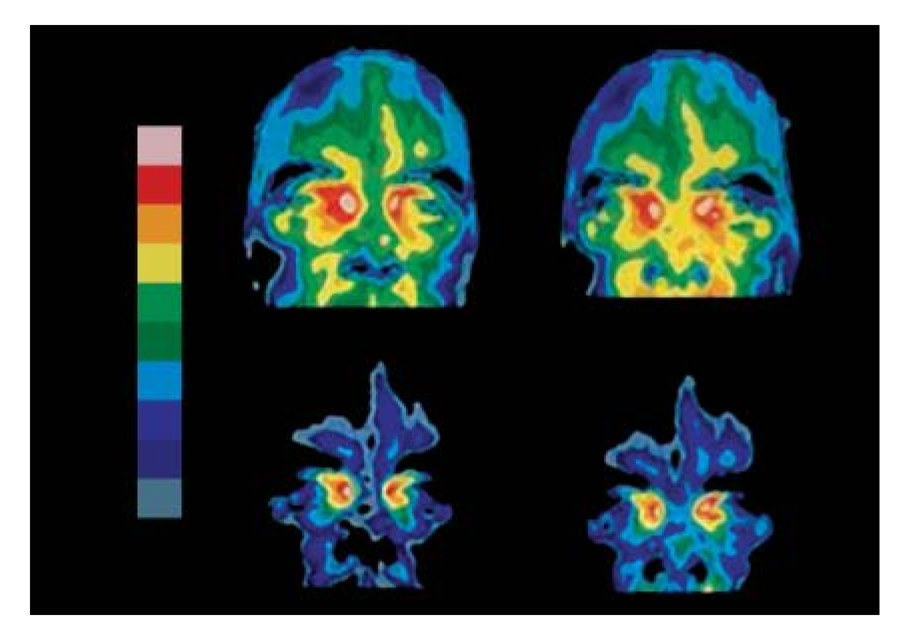
ner el fantoma o animal completo a una radiación de radiofrecuencia de alta intensidad durante un intervalo de tiempo muy corto, se separan rápidamente sus dos partes y se explora la cara interior de una de ellas mediante una cámara termográfica de infrarrojos. Comparando las temperaturas antes y después de la exposición se obtienen las variaciones de temperatura en el plano que interesa. La precisión de esta técnica depende fundamentalmente del tamaño del modelo, ya que la pérdida de calor entre la fase de exposición y la fase de exploración varía mucho con el tamaño del objeto.

La técnica de imágenes por resonancia magnética proporciona una distribución de la temperatura dentro del cuerpo y es la única técnica no invasiva. Sin embargo, su empleo es poco práctico en muchas aplicaciones debido a su elevado costo.

Por último, la técnica de obtención de imágenes por luminiscencia permite obtener mapas de distribución de la tasa de absorción específica en objetos que sean idealmente transparentes. Esta técnica utiliza un compuesto químico-luminiscente, la diazoluminomelanina, y un sistema cuantitativo de imagen luminiscente. El principio físico consiste en el cambio de la intensidad de la luz emitida por el agente químico durante la exposición a la radiación de radiofrecuencia.

SAR promediada

Aparte de estas mediciones locales, cabe medir una SAR promediada en fantomas y cadáveres de animales. Las correspondientes técnicas calorimétricas son muy útiles y fáciles de ejecutar. La más sencilla y económica utiliza un vaso Dewar (una especie de "termo", con paredes de espejo dobles separadas por un vacío). Los procesos de calibración y posibles problemas de operación son muy fáciles de resolver y el tiempo que tiene que transcurrir para obtener un valor de la tasa de absorción específica no resulta muy largo (unos 30 minutos para un ratón). Sin embargo, la técnica calorimétrica más utilizada por los investigadores experimentales es el calorímetro de doble vaso. Para la medida de la tasa de absorción específica se utilizan dos cadáveres de animales de igual



7. LA TECNICA DE LA TERMOGRAFIA se basa en el registro de emisiones térmicas. Esto permite obtener un plano de distribución de temperaturas donde se pueden realizar medidas cuantitativas y cualitativas. En una persona sana, la distribución de cambios de temperatura debe ser relativamente simétrica. Las imágenes muestran la evolución del estrés en una persona en un intervalo de 5 minutos (*imagen de la derecha*).

peso y en equilibrio térmico. Inmediatamente después de exponer uno de los cadáveres a una radiación de RF, se introduce cada uno de ellos en un vaso del calorímetro. El valor de SAR promediado sobre el cuerpo completo del animal se determina a partir de la diferencia en el calor transferido a una cierta cantidad de agua del calorímetro.

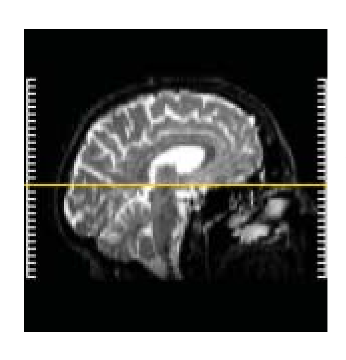
El calorímetro de doble vaso resulta práctico sobre todo con muestras biológicas pequeñas. Sin embargo, su empleo con animales o fantomas grandes presenta grandes dificultades tanto en lo que se refiere al equipo como en lo tocante a la exactitud de la medida. En animales grandes, se invierte mucho tiempo en realizar una medida (aproximadamente un día para una rata de tamaño mediano), y pueden ser necesarios varios días hasta que se alcance la temperatura de equilibrio, debido al largo tiempo de difusión del calor. Durante este tiempo la descomposición del cuerpo genera un calor adicional que introduce errores en el valor de SAR.

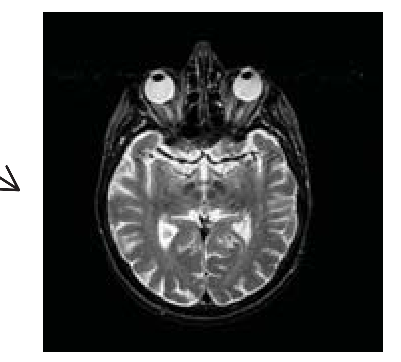
Limitaciones

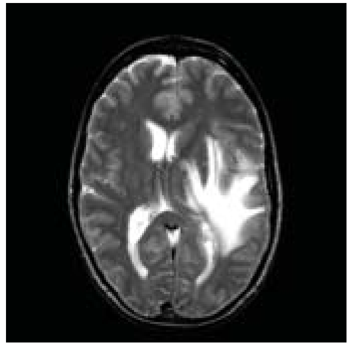
En las secciones anteriores ya se ha mencionado que la SAR ha sido admitida en todo el mundo como una magnitud dosimétrica y en muchos países es el parámetro básico de sus estándares y directivas de seguridad. Así, el estándar ANSI/IEEE C95.1 de niveles de seguridad de los campos de RF en lo que se refiere a la exposición humana, especifica unos valores de SAR para el cuerpo completo en trabajadores de riesgo y población en general de 0,08 y 0,4 watt por kilogramo respectivamente y un valor de SAR local de 8 watt por kilogramo en 1 g de tejido con forma de cubo. Estos valores se han obtenido a partir de estudios sobre la modificación del comportamiento de animales expuestos a una radiación en condiciones de campo lejano, es decir, cuando están alejados de la fuente de radiación. Sin embargo, en este estándar, esos valores máximos de SAR permisibles están especificados sólo en el intervalo de frecuencias comprendido entre 0,1 MHz y 6 GHz. Es evidente que el valor de SAR y su aplicación al establecimiento de estándares se limitan así a un estrecho intervalo de frecuencias. Sería conveniente extenderlos a todo el intervalo de 0 a 300 GHz.

A frecuencias inferiores a los 0,1 MHz, las corrientes inducidas y las corrientes de contacto (que pueden ocasionar *shocks* y quemaduras) cumplen una función impor-

JOSE LUIS SEBASTIAN FRANCO, SAGRARIO MUÑOZ SAN-MARTIN, MIGUEL SANCHO RUIZ Y JOSE MIGUEL MIRANDA PANTOJA







8. IMAGENES POR RESONANCIA MAGNETICA de una sección transversal del cerebro. Cuando los protones del cerebro están inmersos en un campo magnético, transmiten y reciben energía electromagnética. La intensidad de la energía transmitida es proporcional al número de protones en el tejido y viene modificada por las propiedades del microentorno de cada protón, como su movilidad y la homogeneidad local del campo magnético. La señal de la resonancia magnética se puede ponderar para destacar unas zonas de otras, como se muestra en la imagen inferior donde la zona brillante corresponde a un tumor. De la información obtenida, se pueden deducir también valores de la conductividad en las diferentes regiones.

JOSE LUIS SEBASTIAN FRANCO, SAGRARIO MUÑOZ SAN-MARTIN, MIGUEL SANCHO RUIZ Y JOSE MIGUEL MIRANDA PANTOJA

tante, por lo que hay que tenerlas en cuenta, además de las medidas espaciales de los campos eléctrico y magnético. En este rango de bajas frecuencias, se determinan los límites a partir de experimentos de electroestimulación. Por consiguiente, para que los límites de exposición a frecuencias inferiores a 0,1 MHz se puedan establecer según los criterios de SAR o de energía depositada, es fundamental conocer mejor los umbrales dosimétricos tanto para la electroestimulación como para la energía depositada en el intervalo de 0,01 a 1 MHz.

A frecuencias superiores a 6 GHz, la radiación de RF penetra muy poco en el organismo (menos de 2 mm); en este caso la absorción superficial condiciona los posibles peligros y efectos biológicos. De nuevo, si se desea seguir utilizando criterios basados en valores de SAR, es importante estudiar detalladamente la transición de la energía depositada en la superficie del organismo hasta una cierta profundidad del mismo. Uno de los principales problemas de este estudio es la determinación correcta de la dependencia de las constantes térmicas del objeto con la frecuencia, ya que los tiempos promedios de exposición utilizados en el estándar C95.1 están basados en esa dependencia. Los resultados de los distintos trabajos experimentales que se han realizado sobre calentamiento de superficies por IR difieren bastante de los valores de las constantes térmicas que se obtienen de estudios matemáticos teóricos. Una dificultad añadida la constituye el flujo sanguíneo, que redistribuye por conducción la energía térmica.

Otra limitación de la dosimetría actual es su aplicación a volúmenes pequeños o exposiciones parciales. Es bien conocido que la exposición de los testículos y de los ojos a una radiación constituye un importante riesgo. Debido a que hay muy pocos trabajos realizados sobre exposición parcial de objetos biológicos, los límites que se proponen en el estándar C95.1 para exposiciones de este tipo son provisionales. De nuevo, la especificación de unos tiempos medios de exposición más precisos podría eliminar los indeseables efectos de la energía a altas frecuencias en exposiciones parciales de corta duración.

Ya hemos comentado la importancia de conocer bien las propiedades eléctricas del objeto biológico, ya que éstas condicionan la distribución de campos dentro del mismo. Aunque en los últimos cinco años se han mejorado sustancialmente las técnicas de medida, es necesario continuar la investigación experimental sobre la determinación de la permitividad y la conductividad eléctricas de los tejidos. Así, por ejemplo, se ha observado que la permitividad y conductividad del hueso vivo son muy similares a las del músculo. Esto implica que hay que utilizar valores más altos que los que se han venido empleando hasta ahora y que los valores de la SAR que se han calculado están sobrestimados.

Sin embargo, las mayores limitaciones que presenta la dosimetría son la complejidad y la dificultad que acompañan siempre a cualquier medida experimental de SAR. Las variaciones de campo que se producen en el interior del objeto biológico hacen que la incertidumbre de las medidas de SAR local con sondas eléctricas o de temperatura se cifren en torno a 2 decibelios (una medida logarítmica de la intensidad relativa de una perturbación). Cuando los valores medidos de la SAR son inferiores a 10 watt por kilogramo, los datos experimentales carecen de fiabilidad; la incertidumbre es grande. Esto es especialmente importante en las medidas realizadas en exposiciones a campos de baja intensidad, como los originados por los terminales de telefonía móvil. En el caso de la SAR promediada sobre el cuerpo completo, la técnica calorimétrica mejora mucho la exactitud.

No hay sistemas calibrados completos de medida del campo eléctrico interno. Los sistemas de que se dispone hoy en día son muy costosos, requieren una gran experiencia en su manejo y en la interpretación de los resultados y necesitan una gran cantidad de equipos auxiliares. Tampoco hay disponibles métodos estándar para calibrar sondas de campo eléctrico o de temperatura con relación a un valor de SAR fijo. Por último, no se han desarrollado fantomas ni

materiales equivalentes estándar; en su defecto, los investigadores fabrican y emplean en cada caso sus propios modelos, lo que dificulta enormemente la reproducibilidad de los experimentos. Un conocimiento más preciso de las propiedades eléctricas de los tejidos permitirá el desarrollo de fantomas más realistas y unas mejores simulaciones teóricas para el cálculo de la SAR.

Aunque hemos centrado nuestro interés en el campo eléctrico, si un efecto biológico en células o tejidos se debiera a la interacción directa con el campo magnético interno, entonces la tasa de absorción sería insuficiente para valorar los posibles efectos. Hay que hacer medidas de los campos internos eléctrico y magnético, aunque la parte de SAR debida a la absorción del campo magnético resulta mucho más difícil de medir e interpretar que la correspondiente al campo eléctrico.

Los autores

José Luis Sebastián Franco, Sagrario Muñoz San-Martín, Miguel Sancho Ruiz y **José Miguel Miranda Pantoja** son profesores del departamento de física aplicada III de la facultad de físicas de la Universidad Complutense de Madrid. Su actividad investigadora está centrada en los campos electromagnéticos de radiofrecuencias y en sus mecanismos de interacción con objetos biológicos.

Bibliografía complementaria

A PRACTICAL GUIDE TO THE DETERMINATION OF HUMAN EXPOSURE TO RADIOFREQUENCY ELECTROMAGNETIC FIELDS. NCRP, Report 119. National Council on Radiation Protection and Measurements; Bethesda, MD, 1994.

ELECTROMAGNETIC ABSORPTION IN THE HUMAN HEAD FROM EXPERIMENTAL 6 GHZ HANDHELD TRANSCEIVERS. O. P. Gandhi Y J. Y. Chen en *IEEE Transactions on Electromagnetic Compatibility*, vol. 37, n.º 4, pág. 547-558; 1995.

ANALYSIS OF THE INFLUENCE OF THE CELL GEOMETRY, ORIENTATION AND CELL PROXIMITY EFFECTS ON THE ELECTRIC FIELD DISTRIBUTION FROM DIRECT RF EXPOSURE. J. L. Sebastián, S. Muñoz, M. Sancho y J. M. Miranda en *Physics in Medicine and Biology* vol. 46, n.º 1, págs. 213-225; 2001.

Historia de la formación de las estrellas

Los distintos procesos de formación de las galaxias, grandes y pequeñas, condicionaron el ritmo del nacimiento estelar

Alan Heavens

Más de nueve mil trillones (9×10^{21}) de estrellas se han formado en el universo observable en los 13.700 de millones de años que lleva existiendo. A pesar de esa profusión de estrellas, el universo fue bastante oscuro durante sus primeros mil millones de años, según los modelos cosmológicos. En esa época opaca contaba con nubes de gas y materia oscura, y poco más; las primeras estrellas no se formaron hasta pasados cientos de millones de años. Una vez que la maquinaria cósmica productora de estrellas se puso en marcha, la fabricación estelar prosiguió a un ritmo prodigioso. Pero, ¿se mantuvo constante a lo largo de la historia del universo? ¿O se trató de una explosión de natalidad?

Esas preguntas se refieren a algo más que meras curiosidades cósmicas. La explosiva natalidad subsecuente a la Segunda Guerra Mundial reflejaba profundos cambios sociales y económicos en los países occidentales. En una suerte de paralelismo, las distintas tasas de formación estelar nos informan de las circunstancias físicas que rodean al nacimiento de las estrellas. A esas "circunstancias físicas" las llamamos galaxias: la frecuencia de la producción de estrellas está íntimamente ligada al proceso de formación galáctica.

Con la ayuda de los telescopios y ordenadores actuales podemos dar respuestas aceptables a esas preguntas. Con mis colaboradores algo he aportado a la determinación de la historia de la formación de estrellas en el universo. Nos hemos apoyado en el análisis de la luz estelar de 96.545 galaxias. Los resultados sorprenden: a primera vista, parecen contradecir algunos de los principios que rigen la configuración de las galaxias. La paradoja encuentra explicación si se presupone que no todas las galaxias siguieron la misma pauta de formación.

El universo fabrica estructuras

Desde que Vesto Slipher se percató en 1917 de que las líneas espectrales de la luz que recibimos de una galaxia están sistemáticamente desplazadas hacia mayores longitudes de onda (más rojas), contamos con pruebas firmes de que el universo se halla en expansión. La naturaleza de esta expansión depende de la cantidad de masa que hay en el universo; de ahí el interés por determinar la densidad de la materia. Las observaciones recientes de explosiones de estrellas lejanas, de la macroestructura del universo y de la radiación residual de la gran explosión —el fondo cósmico de microondas— nos han mostrado que el universo no cuenta con la materia suficiente para detener la expansión. Es más, la mayor parte de



CORTESIA DE LA NASA, *The Hubble Heritage Team* y ADAM REISS, *Instituto de Ciencias del Telescopio Espacial/American Scientist*

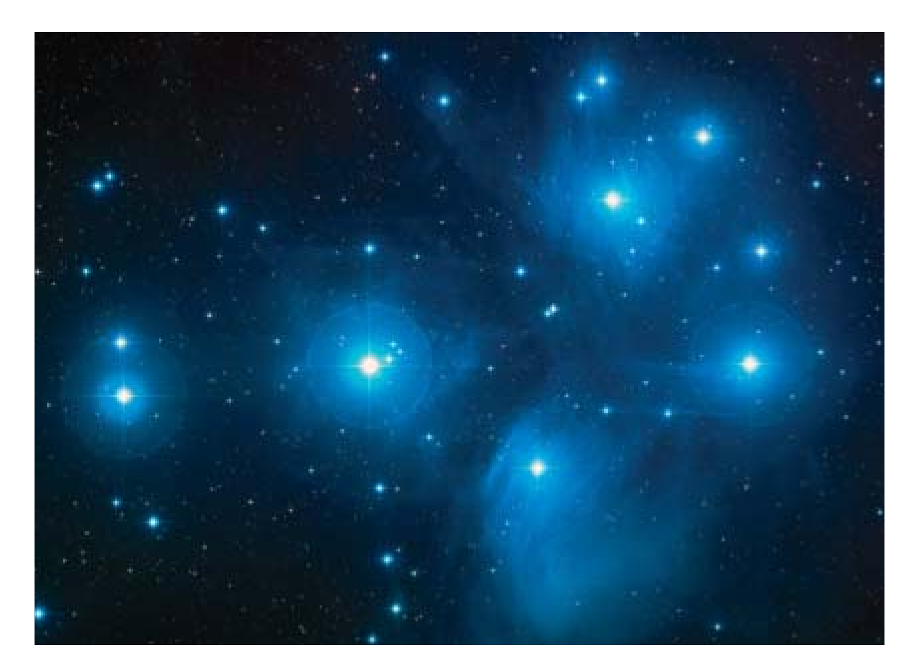
1. LA GALAXIA ESPIRAL NGC 3370 contiene estrellas jóvenes, en las regiones más azules, y una población más envejecida en el centro amarillento. Las diferencias de color indican las temperaturas y edades de las estrellas. El estudio del rango de colores emitido por la galaxia —su espectro— permite determinar la edad relativa de las estrellas y deducir así la tasa de formación estelar en diferentes épocas de la historia galáctica. El promedio de los resultados de muchas galaxias proporciona la historia de la formación de estrellas en el universo.

la materia parece tener una extraña naturaleza. Apenas el 4 por ciento de la densidad consta de neutrones, protones y electrones, la materia de que están hechos los planetas y de la que estamos hechos nosotros. Un 23 por ciento consiste en materia oscura, una forma de materia extraordinaria y desconocida que no se observa en la Tierra. Y el 73 por ciento es energía oscura, una "sustancia" aún más rara: como presenta "gravedad repulsiva", acelera la expansión del universo.

A pesar de los aspectos que siguen envueltos en el misterio, conocemos los mecanismos de formación de las estructuras que apreciamos hoy. Las observaciones del fondo cósmico de microondas, que data de cuando el universo tenía sólo 300.000 años, muestran la falta de uniformidad de éste en su primera época. Las han obtenido diversos instrumentos instalados en tierra, en globos y en el espacio; el más famoso de ellos fue el Explorador del Fondo Cósmico (COBE), de principios de los noventa; ahora destaca la sonda Wilkinson para la Anisotropía de Microondas, de mayor resolución que COBE. Esas mediciones han descubierto unas pequeñas irregularidades en la densidad del universo primigenio. Se trataba de irregularidades inestables: al tener mayor densidad que otras regiones, su gravedad era un poco superior, por lo que atraían hacia sí materia y la acumulaban. Mediante la combinación de una acreción suave y la fusión de pequeñas unidades, se constituyeron, andando el tiempo, los objetos densos.

Aunque parezca paradójico, para que una estrella tan caliente se forme, ha de enfriarse primero el gas. Las nubes grandes de gas se sostienen por presión si están calientes, pero si el gas se enfría, entonces la presión se relaja y la nube colapsa. A las escalas más pequeñas, las fases últimas de ese derrumbe calientan la nube y, si la temperatura alcanza varios millones de grados, comienzan las reacciones nucleares y el objeto se convierte en una estrella.

A mayores escalas, el proceso de colapso sucede de abajo arriba. Primero colapsan objetos no muy grandes (de alrededor de un millón de veces la masa solar). A medida que el gas denso se enfría, se fragmenta en nubes más pequeñas, cuyo derrumbe crea estrellas. Estos cúmu-



2. LAS ESTRELLAS CALIENTES DE GRAN MASA, como las del cúmulo de las Pléyades, emiten luz ultravioleta. Las estrellas de masa imponente, como éstas, tienen una vida corta; la presencia de luz ultravioleta en una galaxia apunta a una formación reciente de estrellas en su seno. El método tradicional de medición de la tasa de formación de estrellas explota esta propiedad y analiza la emisión ultravioleta de las galaxias lejanas.

los de muchas estrellas se agrupan, por efecto de la gravedad, en sistemas mayores, que pueden también mezclarse entre sí. Y de ese modo continúa el proceso. Jerárquicamente, se van construyendo sistemas de tamaño creciente: galaxias pequeñas, galaxias grandes, grupos de galaxias, cúmulos de galaxias y supercúmulos galácticos. Hoy día, las estructuras de mayor tamaño que han sufrido su desplome absoluto son los cúmulos de galaxias. Contienen materia oscura, gas y galaxias repletas de estrellas; en total suman cientos de billones de masas solares.

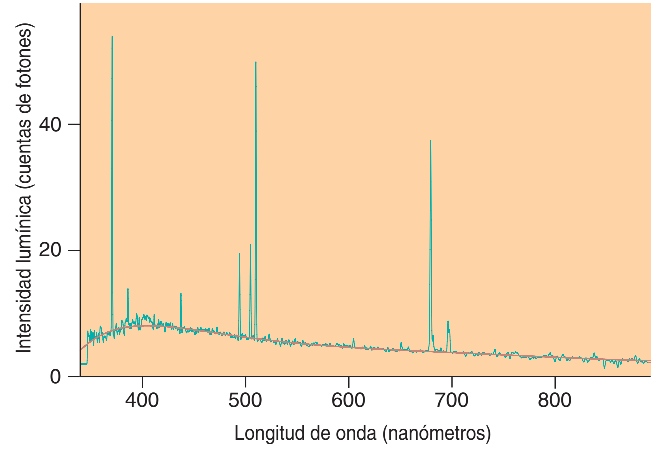
Gracias a este esquema jerárquico de la formación de galaxias conocemos el mecanismo de rotación de las galaxias espirales, por qué hay galaxias de masa muy dispar y por qué existe un límite para los mayores objetos observados. Cierto es que muchos detalles se nos escapan en su pleno sentido. La formación de una galaxia es un proceso complejo; nos hallamos lejos de dominar el proceso a través del cual transforma su gas en estrellas. Así, aunque sabemos que las estrellas influyen en el gas mediante las supernovas y los vientos en un proceso de "retroalimentación", y que la retroalimentación puede favorecer o inhibir la formación estelar, carecemos de estimaciones fiables de los mecanismos involucrados.

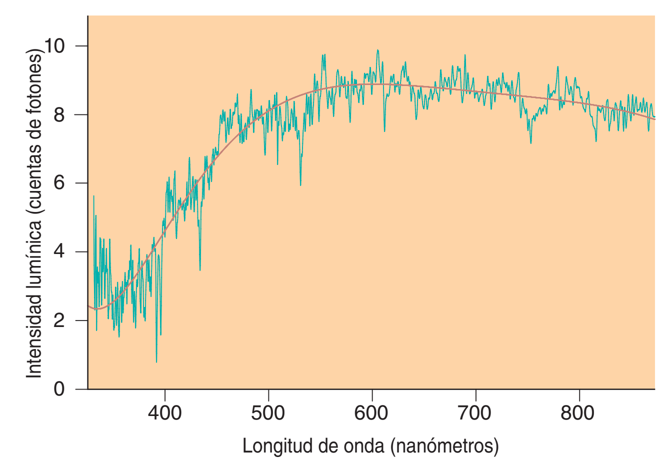
Algunos astrónomos y cosmólogos intentamos determinar el momento en el que el universo creó la mayoría de las estrellas y la trayectoria evolutiva y cambiante de esa tasa de formación.

Fósiles contra recién nacidos

El problema de averiguar cuántas estrellas se han ido formando con el tiempo recuerda al de establecer la historia de la tasa de natalidad humana. Se podría indagar en los archivos históricos del registro civil o sacar instantáneas de personas y estimar su edad. Si se utiliza el segundo método se deberían incluir todos los lugares donde haya seres humanos, tumbas incluidas. En el primer método importa disponer de unos registros históricos exhaustivos. Si no hay errores, los dos métodos deberían producir resultados análogos.

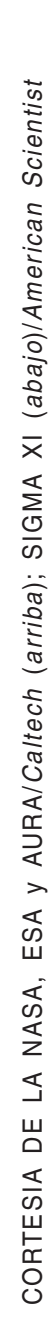
De igual forma, la historia de la tasa de formación estelar puede determinarse de dos maneras. El método tradicional se basa en la velocidad finita de la luz, que nos permite ver tiempos pasados. Al analizar la luz



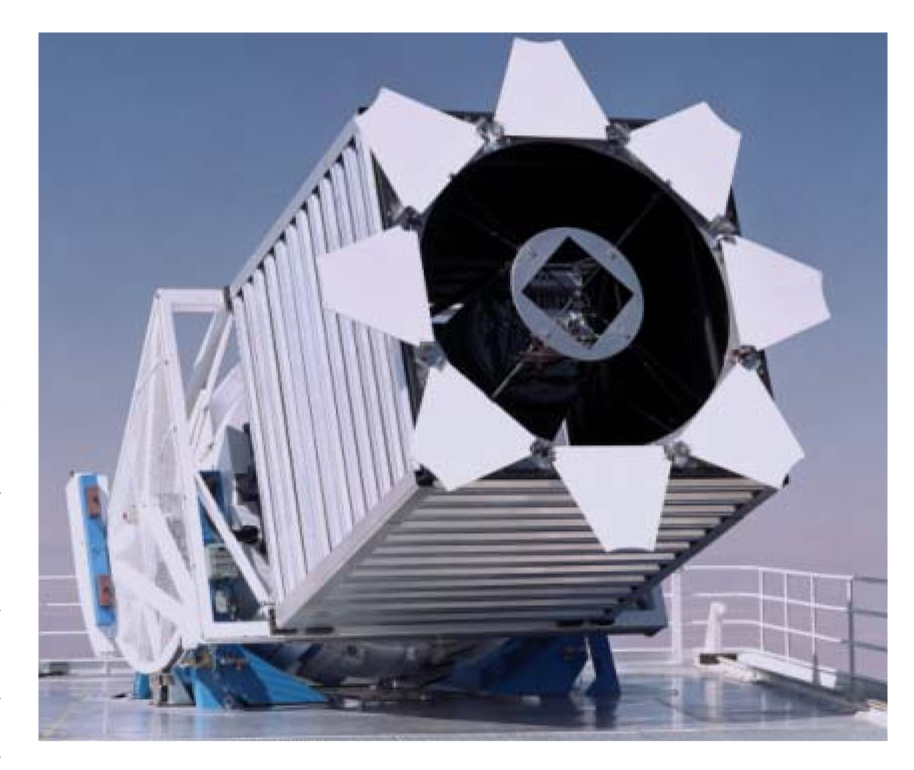


3. EL ESPECTRO DE UNA GALAXIA ilustra las cantidades relativas de luz emitidas a distintas longitudes de onda. Una galaxia con una población joven de estrellas emite más luz en la parte azul del espectro, alrededor de los 400 nanómetros (*izquierda*).

Por el contrario, una galaxia con una población estelar más antigua emite sobre todo luz en el rojo, por encima de los 500 nanómetros (*derecha*). El nuevo método para medir la tasa de formación estelar hace uso del espectro completo de una galaxia.



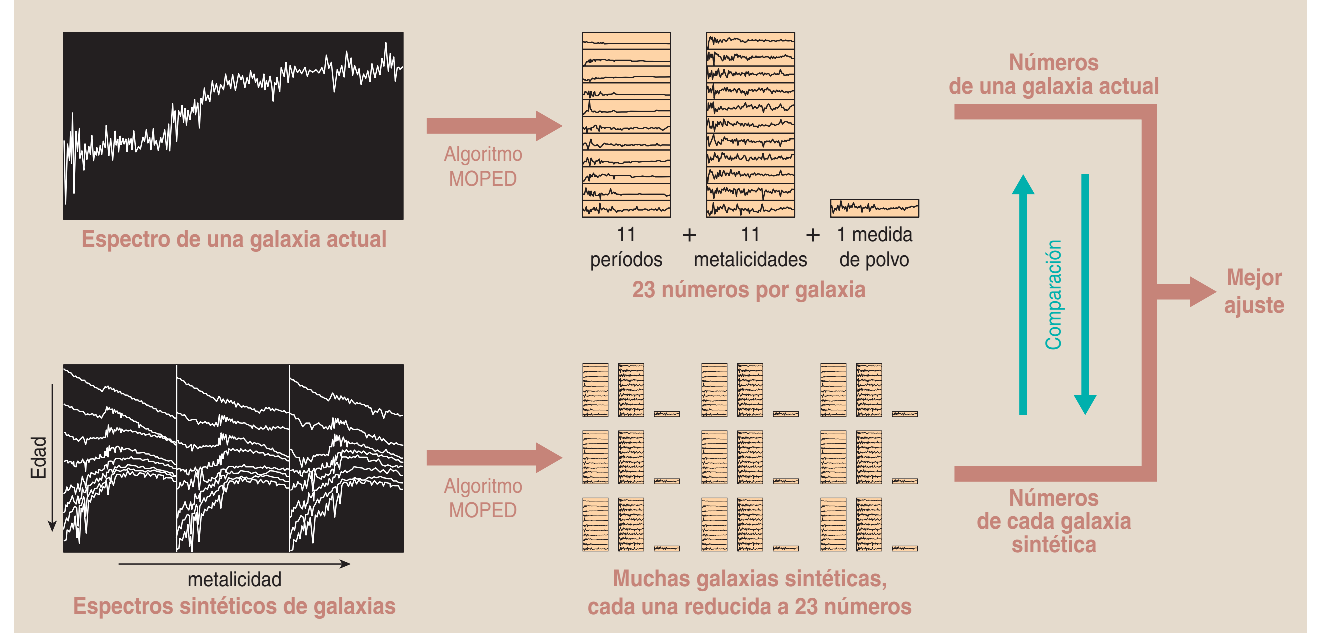
4. EL ESTUDIO DIGITAL SLOAN recorre los cielos con un telescopio de abertura de 2,5 metros. Los dos espectrógrafos del telescopio recolectan la luz de miles de galaxias cada noche. Se han analizado los espectros de casi 100.000 galaxias para determinar la tasa de formación estelar del universo.



procedente de objetos muy lejanos, los vemos tal y como eran hace miles de millones de años. Se buscan indicios de estrellas recién formadas por entonces. El método alternativo adopta un enfoque "fósil". Se estudia una muestra muy amplia de galaxias cercanas y se calculan las edades de las estrellas de cada una. Este es el enfoque que he adoptado, en estrecha colaboración con Raúl Jiménez, de la Universidad de Pennsylvania, y Ben Panter y James Dunlop, de la Universidad de Edimburgo.

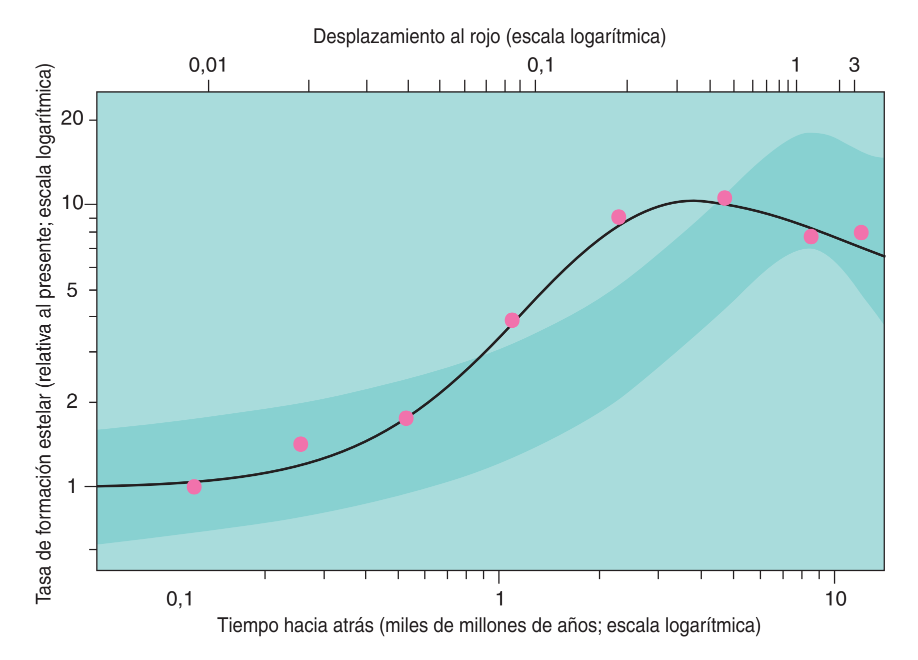
Hasta hace poco, el análisis de la historia de la formación estelar del universo ha estado dominado por la búsqueda de signos que revelasen la presencia de estrellas recién nacidas; en la analogía, vendría a ser la lectura de los "registros históricos" del nacimiento estelar. Con frecuencia se observaban galaxias lejanas, es decir, con un notable corrimiento hacia el rojo.

Un indicador de la formación reciente de estrellas es la cantidad de luz ultravioleta emitida. La luz ultravioleta procede sobre todo de estrellas muy calientes, como las que se observan en el cúmulo estelar de las Pléyades. Las estrellas muy calientes poseen una masa imponente, queman su combustible nuclear muy deprisa y mueren pronto; por tanto, sólo las vemos jóvenes. La presencia de luz ultravioleta es sin duda un



5. EL ANALISIS DEL ESPECTRO DE UNA GALAXIA ha de compararse con los espectros sintéticos de muchas galaxias hasta encontrar el mejor ajuste. Las galaxias sintéticas divergen en función de su edad y composición química ("metalicidad"). El algoritmo MOPED reduce el espectro de cada galaxia a 23 números, el resultado de la suma ponderada de cada espectro. El mejor ajuste de los 23 números de la galaxia real con los de las galaxias sintéticas proporciona la historia de la formación estelar, la de la metalicidad y el contenido de polvo de la galaxia.

CORTESIA DE FERMILAB VISUAL MEDIA SERVICES (*arriba*); SIGMA XI (*abajo*)/*American Scientist*



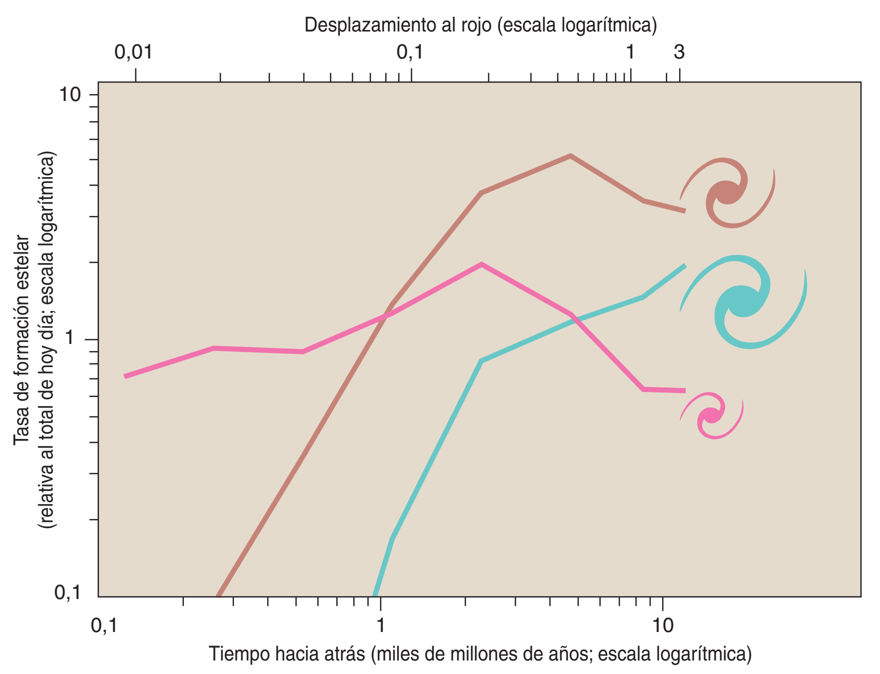
6. EL PROMEDIO DE LA TASA DE FORMACION estelar ha cambiado con el tiempo. El universo tiene unos 13.700 millones de años, pero la tasa de formación de estrellas llegó a su máximo hace unos 5000 millones (*línea negra*). Las determinaciones anteriores de la tasa de formación (*la banda de color azul*) apuntaban a una actividad máxima más precoz, hace unos 8000 millones de años. Esta discrepancia se explica por el sesgo que los métodos tradicionales introducen en la población de galaxias que estudian. Aquí se muestran ocho de los once períodos temporales (*puntos rosados*) de las historias galácticas.

buen indicador de estrellas recién formadas en una galaxia.

Por otra parte, a medida que la población estelar envejece, las estrellas azules enrojecen y la luz estelar adopta un tono rojizo o amarillento. Existen otros indicadores de una formación reciente de estrellas, como la emisión de líneas de hidrógeno procedentes del gas ionizado que rodea las estrellas masivas. La luz azul de estos astros ioniza enormes volúmenes de hidrógeno y los hace brillar. Sin embargo, el estudio de galaxias con un corrimiento hacia el rojo consume mucho tiempo, además, amén de hallarse limitado a la observación de los sistemas de mayor masa porque los pequeños resultan demasiado débiles para detectarlos a grandes distancias.

Muy distinto es el enfoque fósil. Se basa en el espectro de la luz de una galaxia a lo largo de un rango amplio de longitudes de onda. La mayor parte de la luz procede de las estrellas que componen la galaxia, aunque el gas caliente emite una fracción. Si eliminamos las líneas de emisión del gas, nos quedamos sobre todo con luz estelar. Cada estrella contribuye de una manera singular al espectro, en razón de su edad y de la composición química del gas del que se formó. Puesto que los elementos se fabrican dentro de las estrellas, los elementos más pesados —llamados "metales" por los astrónomos— se acumulan en el gas y en las nuevas generaciones estelares a medida que la galaxia evoluciona. El estudio minucioso del espectro nos proporciona la edad de las estrellas y revela cuántas de ellas se formaron en diferentes períodos de la historia galáctica. Podemos entonces deducir el cambio de la composición química del gas con el tiempo.

Consideremos el espectro de una galaxia que alberga estrellas jóvenes, incluidas algunas dotadas de masa enorme que emiten luz azul. Esta característica se manifiesta a través de emisiones intensas en longitudes de onda "azules", alrededor de los 400 nanómetros. Y sea el espectro de otra galaxia que presente muy poca luz azul, indicio de que las estrellas de gran masa y calientes ya no brillan. Una galaxia así ha tenido poca formación estelar



7. LA TASA DE FORMACION ESTELAR varía con la masa de la galaxia. Las galaxias mayores (*azul*) alcanzan la tasa máxima antes que las galaxias de tamaño intermedio (*marrón*) y las pequeñas (*rosado*). Los métodos tradicionales para determinar la tasa de formación de estrellas en el universo presentan un sesgo hacia las galaxias mayores y conducen, por ello, a un resultado equivocado.

SIGMA XI/*American Scientist*

8. LOS DIFERENTES PROCESOS DE FORMACION de galaxias grandes y pequeñas afectan al ritmo de producción de estrellas. Una galaxia muy pequeña puede crearse a partir de una nube solitaria, de tamaño moderado, que empezó a fabricar estrellas cuando el universo llevaba existiendo ya un tiempo considerable. Por el contrario, una galaxia grande pudo originarse por la mezcla (fusión) de muchas nubes pequeñas de gas que crearon sus estrellas en una época precoz de la historia del universo. No se contradice el principio general que establece que en el universo las estructuras menores tienden a desarrollarse antes que las mayores. Las fusiones destrozan la estructura espiral; las galaxias mayores suelen ser elípticas.

SIGMA XI/American Scientist

reciente; la población es vieja, la luz está dominada por las estrellas rojas y frías. Tras el análisis de un número sustantivo de estos espectros llegamos a vislumbrar cuánta formación estelar ha acontecido en un momento particular de la historia del universo.

MOPED

El Estudio Digital Sloan de los Cielos es una fuente de especial valor para nuestro trabajo. Cubrirá cerca de un cuarto del cielo y medirá el espectro de cientos de miles de galaxias hasta una distancia de tres mil millones de años-luz. La parte norteamericana del proyecto emplea un telescopio de 2,5 metros instalado en Punta Apache, Nuevo México. Está conectado a 30 detectores de carga acoplada (CCD) que producen cada noche 200 gigabytes de datos de imágenes. En cada instantánea se obtienen los espectros de 640 galaxias, cuya luz llega, a través de fibras ópticas, a dos espectrógrafos.

No es fácil interpretar un espectro. Por comodidad de cálculo, dividimos la historia de una galaxia en 11 períodos. Determinamos la fracción de estrellas de la galaxia nacidas en cada período. También analizamos la composición química galáctica para cada intervalo temporal. La tarea se complica porque muchas galaxias contienen polvo que absorbe la luz azul, por lo que la galaxia aparece más roja de lo que realmente es. En total, los 11 períodos de formación estelar, las 11 mediciones de la composición química y una del polvo suman 23 números por galaxia.

Para obtener esos valores buscaremos un espectro modelo que coincida en el mayor grado posible con el observado. En teoría resulta posible, pero en la práctica se hace exasperantemente lento, sobre todo considerando que tenemos cerca de cien mil galaxias para analizar. Por eso recurrimos a un algoritmo de compresión de datos —MOPED (siglas en inglés de Compresión y Estimación Múltiple Optimizada de Datos)— que reduce cada galaxia a 23 números en sustitución de los varios miles de puntos del espectro. El análisis centuplica así su velocidad sin merma de precisión. MOPED es una técnica patentada que se concibió especialmente para nuestros objetivos, pero es aplicable a otros problemas de búsqueda inversa muy dispares. No obstante, se necesitan varias semanas para que 20 computadoras procesen el sondeo galáctico completo.

MOPED nos proporciona una historia completa de la formación estelar para cada galaxia. Si sumamos adecuadamente todas las galaxias de la exploración, tendremos el promedio de la tasa de formación estelar

por unidad de volumen en el universo. Nuestros resultados concuerdan con los ya conocidos por otras técnicas para la época primigenia del universo y para su historia más reciente, pero no coinciden en los tiempos intermedios. Según los métodos tradicionales, la tasa de formación de estrellas llegó a su máximo hace unos 8000 millones de años; los nuevos, en cambio, indican que el máximo ocurrió hace 5000 millones, más o menos cuando se formaron el Sol y el sistema solar. Los dos métodos revelan que la formación estelar ha decaído rápidamente; hoy en día la tasa es apenas entre un 10 y un 15 por ciento de lo que fue en su apogeo. El nuevo máximo encontrado por el "método del fósil" debería confirmarse por otros métodos, pero no será fácil, dada la dificultad de observar todas las galaxias a grandes distancias y no sólo las más brillantes.

¿Cómo explicar una discrepancia de nada menos 3000 millones de años? Para lograrlo, acudiremos a las tasas de formación estelar de galaxias de masas diferentes. Hay una tendencia obvia: las galaxias de mayor masa presentan antes el máximo de formación estelar. Es decir, la mayoría de las estrellas de las galaxias grandes, nuestra Vía Láctea entre ellas, nacieron pronto; ahora no nacen muchas estrellas en su seno. Nuestra galaxia apenas produce ya un puñado de estrellas nuevas por año. Por otro lado, las galaxias de poca masa forman estrellas más tarde y continúan creándolas a un buen ritmo.

La clave se encuentra en que los métodos tradicionales, fundamentados en la observación de las galaxias lejanas, sólo detectan las que gozan de mayor masa. Tal y como hemos aprendido en la exploración Sloan, las mayores galaxias no son representativas porque han formado sus estrellas mucho antes. Por tanto, no sorprende que las determinaciones anteriores encontraran un máximo temprano de la tasa de formación estelar. Siguiendo con la analogía de la natalidad humana, es como si se hubieran perdido los archivos de las parroquias y hubiera que utilizar sólo los de las catedrales. Valdría mientras las catedrales fueran representativas del conjunto, pero si los cambios demográficos hicieron que archivasen una fracción cada vez menor de los nacimientos, sus archivos nos estarían proporcionando una visión sesgada de la realidad.

Conclusión

¿Cómo encaja todo esto con las teorías sobre la formación de las galaxias? De entrada pone en cuestión el modelo jerárquico, según el cual la formación de los objetos pequeños precede a la de los objetos grandes. Observamos lo contrario: las galaxias de mayor masa producen estrellas antes que las galaxias menores. Pero no debemos olvidar un aspecto importante de nuestro escrutinio de las galaxias del Estudio Sloan. Analizamos las estrellas tal y como las vemos hoy día y deducimos cuándo nacieron. Pero no tenemos un conocimiento directo de dónde se formaron. En concreto, no existe razón alguna para pensar que las estrellas nacidas hace ocho mil millones de años se originaron en la galaxia donde las vemos. Quizá se crearon en galaxias más pequeñas que luego se mezclaron con otras en algún momento de su historia. Y es muy probable que haya ocurrido así: la mayoría de las estrellas que componen una galaxia grande nacieron casi con toda certeza en unidades menores que luego se fueron ensamblando hasta crear la gran galaxia.

Las galaxias de baja masa siguen probablemente un proceso parecido, pero al revés. Primero se reúne toda la masa, y ésta atrae al gas de los alrededores. El gas que cae alimenta la formación de estrellas, o incluso desencadena la producción estelar en el gas ya residente en la galaxia. Se podría ir más lejos y afirmar que en el modelo jerárquico de la formación galáctica de abajo arriba, los objetos que terminan con mayor masa son los que han producido sus estrellas precozmente.

Una de las consecuencias más atractivas de disponer de dos métodos para el estudio de estas cuestiones es la posibilidad que ofrece de comprobar el principio copernicano. El principio establece que nosotros no ocupamos un lugar especial en el universo. Puesto que MOPED nos proporciona información sobre la tasa de formación estelar aquí en el pasado y los otros métodos nos facilitan la tasa de formación a distancias muy lejanas, podemos comparar los dos resultados y comprobar si nuestra posición es atípica. Puesto que creemos entender el porqué de las diferencias entre los dos métodos de trabajo, concluimos que el principio de Copérnico es cierto a este respecto.

Sabemos que nuestro universo experimentó una explosión demográfica estelar hace unos 5000 millones de años. ¿Es posible otra explosión más adelante? Hemos dejado atrás el apogeo de la formación de estrellas; la tasa se está decelerando. A medida que se constituyen estructuras mayores en el universo, el gas se va calentando, lo que dificulta que se enfríe y colapse para producir nuevas estrellas. Además, el gas del universo se enrarece con el tiempo; es decir, se está agotando el suministro necesario para fabricar estrellas. En el futuro, los cielos se oscurecerán. Pero no es algo que debamos temer: el Sol tiene por delante otros cinco mil millones de años.

El autor

Alan Heavens es profesor de astrofísica teórica de la Universidad de Edimburgo. Estudia la macroestructura del universo, la formación de galaxias y las lentes gravitatorias.

Bibliografía complementaria

THE STAR-FORMATION HISTORY OF THE UNIVERSE FROM THE STELLAR POPULATIONS OF NEARBY GALAXIES. A. F. Heavens, B. D. Panter, R. Jiménez y J. S. Dunlop en *Nature*, vol. 428, págs. 625-627; 2004.

SYNTHETIC STELLAR POPULATIONS: SINGLE STELLAR POPULATIONS, STELLAR INTERIOR MODELS AND PRIMORDIAL PROTO-GALAXIES. R. Jiménez, J. MacDonald, J. Dunlop, P. Padoan y J. Peacock en *Monthly Notices of the Royal Astronomical Society*, vol. 349, págs. 240-254; 2004.

THE STAR FORMATION RATE OF THE UNIVERSE AT Z ∼ 6 FROM THE HUBBLE ULTRA-DEEP FIELD. A. J. Bunker, E. R. Stanway, R. S. Ellis y R. G. McMahon en *Monthly Notices of the Royal Astronomical Society*, vol. 355, págs. 374-384; 2004.

Acceso inalámbrico a la Red

El acceso sin cable a la Red se utiliza cada vez más. Esta técnica se está perfeccionando para mejorar su velocidad y fiabilidad

Alex Hills

Son muchos los partidarios de entrar en Internet sin cable. Dicho con la expresión al uso, acceso Wi-Fi ("Wireless Fidelity"). La conexión inalámbrica se utiliza en cibercafés, salas de espera de aeropuertos y en el hogar. El sistema Wi-Fi resulta atractivo porque pone la Red a disposición de los usuarios en todo momento y en cualquier lugar. Proporciona conexiones rápidas que permiten la recepción casi instantánea de mensajes de correo electrónico y una rauda presentación en pantalla de páginas de la Red. Y todo ello con la movilidad y libertad que ha hecho casi ubicua a la telefonía móvil.

Una firma que investiga en el sector de las comunicaciones, Pyramid Research, pronostica que el número total de usuarios Wi-Fi en todo el mundo podría alcanzar los 271 millones en 2008, con 177 millones sólo en Estados Unidos. En la actualidad, la comunidad de usuarios de Wi-Fi sostiene ya una pujante industria internacional que factura unos 2500 millones de euros anuales, según extrapolaciones de cifras obtenidas por In-Stat, otra compañía de prospección de mercados. Pero la misma popularidad de la técnica Wi-Fi es también causa de problemas. Conforme se utilicen con mayor intensidad, las redes Wi-Fi pueden verse incapaces de habérselas con el aumento de tráfico; como resultado, los equipos de los clientes pueden bloquearse o sufrir largas esperas o lentitud en el servicio.

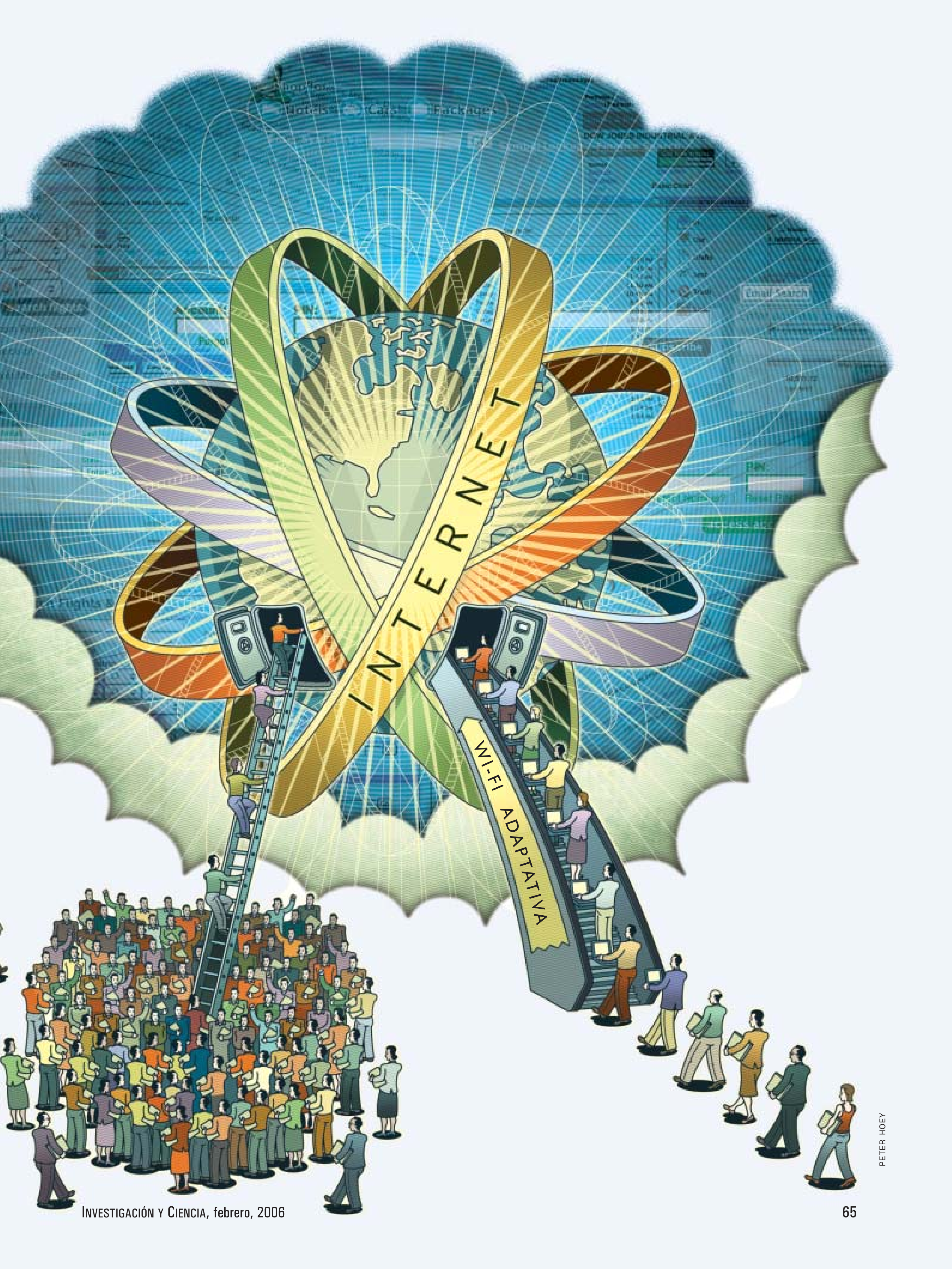
El acceso inalámbrico no logra, ni siquiera funcionando a la perfección, alcanzar la velocidad de las conexiones alámbricas a Internet, como las líneas ADSL o de cable, por ejemplo. No es de esperar que las señales de radio alcancen las velocidades de transmisión que consienten los cables de cobre o los de fibra óptica. Tampoco pueden la Wi-Fi ni otras técnicas inalámbricas basadas en radiodifusión garantizar la seguridad de los otros sistemas, pues la transmisión puede interceptarse mediante receptores cercanos.

Muchas de estas deficiencias quedaron manifiestas ya en 1993, cuando el autor dirigía un equipo de la Universidad Carnegie-Mellon que iba a construir Wireless Andrew, la primera red de área local inalámbrica a gran escala y precursora de las actuales redes Wi-Fi. El proyecto se finalizó en 1999. En la actualidad, Wireless Andrew sustenta la conexión de las instalaciones universitarias.

Mucho ha cambiado el mundo de lo inalámbrico en los doce años transcurridos desde la concepción de nuestra red inalámbrica en Carnegie-Mellon. El notable incremento del uso de la técnica Wi-Fi ha generado varios problemas. Pero, al propio tiempo, se han dado pasos importantes para resolverlos. Antes de examinar estos avances, conviene aclarar cómo opera el sistema Wi-Fi.



1. EL ACCESO WI-FI mejorará en cuanto las redes inalámbricas consigan habérselas con el creciente número de usuarios de equipos móviles que tratan de entrar en Internet.



PETER HOEY

Las entrañas de Wi-Fi

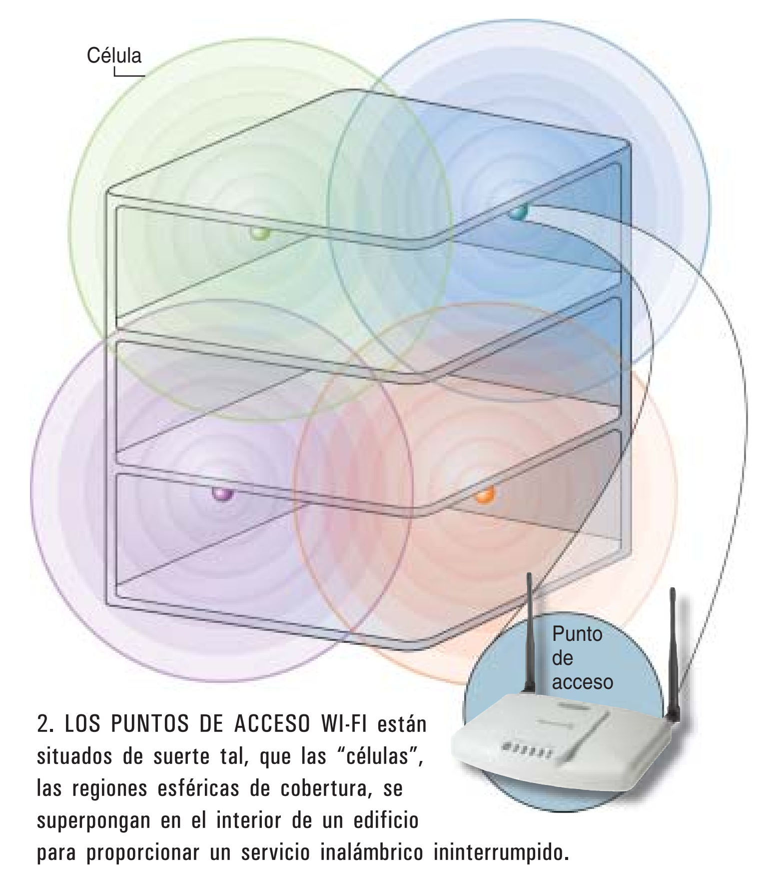
Las redes Wi-Fi comprenden ordenadores móviles equipados con Wi-Fi (sean portátiles o de mano) o terminales especiales de telefonía Wi-Fi, así como los puntos de acceso (AP). Los AP corresponden a estaciones base que se comunican por radio y por cable con los sistemas móviles, así como con las redes que en última instancia dan acceso a Internet. Cada AP envía y recibe señales en un ámbito limitado que suele oscilar entre los 20 y 50 metros en el interior de un edificio. La región de cobertura de un punto de acceso forma una célula esférica (similar a la de un teléfono móvil, aunque sensiblemente menor) que da servicio simultáneo a un gran número de dispositivos móviles alojados en su interior.

Al principio, las redes Wi-Fi se denominaron redes locales inalámbricas (LAN, de "local-area network"). Antes de 1997, los equipos para una LAN inalámbrica no eran interoperativos (los sistemas de un fabricante no se "hablaban" con los de otras compañías). Pero en 1997, el Instituto de Ingeniería Eléctrica y Electrónica (IEEE) adoptó el estándar IEEE 802.11, que eliminaba la incompatibilidad.

En la actualidad, casi todos los equipos de LAN inalámbrica se ajustan al estándar, cuya denominación popular es Wi-Fi (fidelidad inalámbrica). Aunque la norma no dicta todos los aspectos del funcionamiento de la red, sí asegura que equipos de diferente tipo puedan operar de forma conjunta.

Cuatro son los principales retos del desarrollo de las redes Wi-Fi: garantizar la fiabilidad (asegurar que el servicio no se verá interrumpido por radiotransmisiones de mala calidad), mantener las prestaciones (evitar conexiones lentas y demoras excesivas), ofrecer una cobertura total y proporcionar seguridad frente a usuarios no autorizados o escuchas inalámbricas hostiles.

Tales problemas se deben, sobre todo, a los inconvenientes de la radiodifusión, el sistema de transmisión que utilizan las redes LAN inalámbricas. La señal recibida por un cliente o un punto de acceso puede verse degradada de varias formas:



2. LOS PUNTOS DE ACCESO WI-FI están situados de suerte tal, que las "células", las regiones esféricas de cobertura, se superpongan en el interior de un edificio para proporcionar un servicio inalámbrico ininterrumpido.

- Una emisión inalámbrica sufre atenuación (se debilita) al aumentar la distancia, aun cuando no existan obstáculos; si los hubiera, mermarían todavía más la intensidad de la señal.
- Las ondas de radio pueden sufrir una distorsión multitrayecto, provocada por la reflexión en paredes, edificios, mobiliario, electrodomésticos u otros objetos cercanos. Las señales viajan entonces desde el emisor hasta el receptor por diversas vías; llegan a éste varias copias de una misma transmisión, cada una en un instante ligeramente distinto. Los duplicados tardíos corrompen la señal directa (la que viaja sin distorsión), creando problemas en la recepción.
- Un tercer tipo de degradación de la señal se debe al ruido radioeléctrico y a las interferencias. Las interferencias resultan de la superposición de dos o más emisiones de radio. Entre las fuentes de interferencias en redes Wi-Fi se cuentan los hornos de microondas, que radian señales parásitas. Felizmente, los hornos de microondas modernos están dotados de una protección que minimiza tales emisiones. Es cierto que la naturaleza engendra ruido radioeléctrico, pero casi todo el existente procede de maquinaria eléctrica, motores de automóvil, luces fluorescentes u otras fuentes de factura humana.

Los ingenieros de comunicaciones saben cómo vencer estas dificultades, pero sus métodos ralentizan la velocidad de transmisión. Mientras que las redes Ethernet conectadas mediante cables eléctricos ofrecen velocidades de 100 a 1000 megabit por segundo (Mbps), muchas LAN inalámbricas se valen del estándar IEEE 802.11b y,

Resumen/Redes locales inalámbricas

- Al generalizarse el uso de la técnica Wi-Fi (acceso inalámbrico a Internet), el aumento de tráfico amenaza con saturar las redes locales (LAN), basadas en conexiones por radio, que se utilizan para entrar en la Telaraña Mundial. Los usuarios sufrirían entonces retrasos inaceptables e incluso la interrupción del servicio. La Wi-Fi de segunda generación, o adaptativa, y otras mejoras técnicas avanzan en la resolución de estas deficiencias.
- Cuatro son los actuales puntos débiles de las redes inalámbricas: pérdida de calidad en las transmisiones de radio, lentitud de las conexiones (con las consiguientes demoras), cobertura insuficiente e inseguridad de las comunicaciones. La Wi-Fi adaptativa, que está empezando a entrar en servicio, resolverá esas deficiencias y ofrecerá prestaciones adicionales.

LUCY READING-IKKANDA (*ilustración*); SYMBOL TECHNOLOGIES, INC. (*router inhalámbrico*)

por tanto, operan a una velocidad máxima de 11 Mbps. Equipos más modernos, basados en IEEE 802.11a y 802.11g, alcanzan hasta 54 Mpbs (todavía unas "tortugas" al lado de las "liebres" de Ethernet). Está próxima la introducción de una versión de IEEE 802.11 que permitirá comunicaciones a 108 Mbps.

En realidad, los valores mencionados exageran la velocidad real de transmisión de los sistemas Wi-Fi, pues ésta se reduce de forma automática desde su valor máximo (sea de 11 o de 54 Mbps) a uno inferior por atenuación de la señal, recepción multitrayecto, ruido o interferencias. Una conexión IEEE 802.11b puede reducir la velocidad de 11 Mbps a 5,5 Mbps, a 2, e incluso a sólo 1 Mbps. Además, los bits de redundancia (bits extra que se añaden a cada transmisión para controlar el funcionamiento de la red y minimizar errores) reducen más todavía la tasa efectiva de transferencia de datos.

Desde la introducción de la Wi-Fi inicial, el autor y sus compañeros en Carnegie-Mellon y Airespace (ahora integrada en Cisco Systems), así como ingenieros de otras universidades y compañías, vienen trabajando para mejorar su fiabilidad, rendimiento, diseño y seguridad. El resultado, la segunda generación de equipos Wi-Fi ("Smart Wi-Fi", o Wi-Fi adaptativa), incorpora varias prestaciones nuevas destinadas a reparar las deficiencias de la técnica. Estas ampliaciones se fundan en una mayor "inteligencia" de los sistemas Wi-Fi.

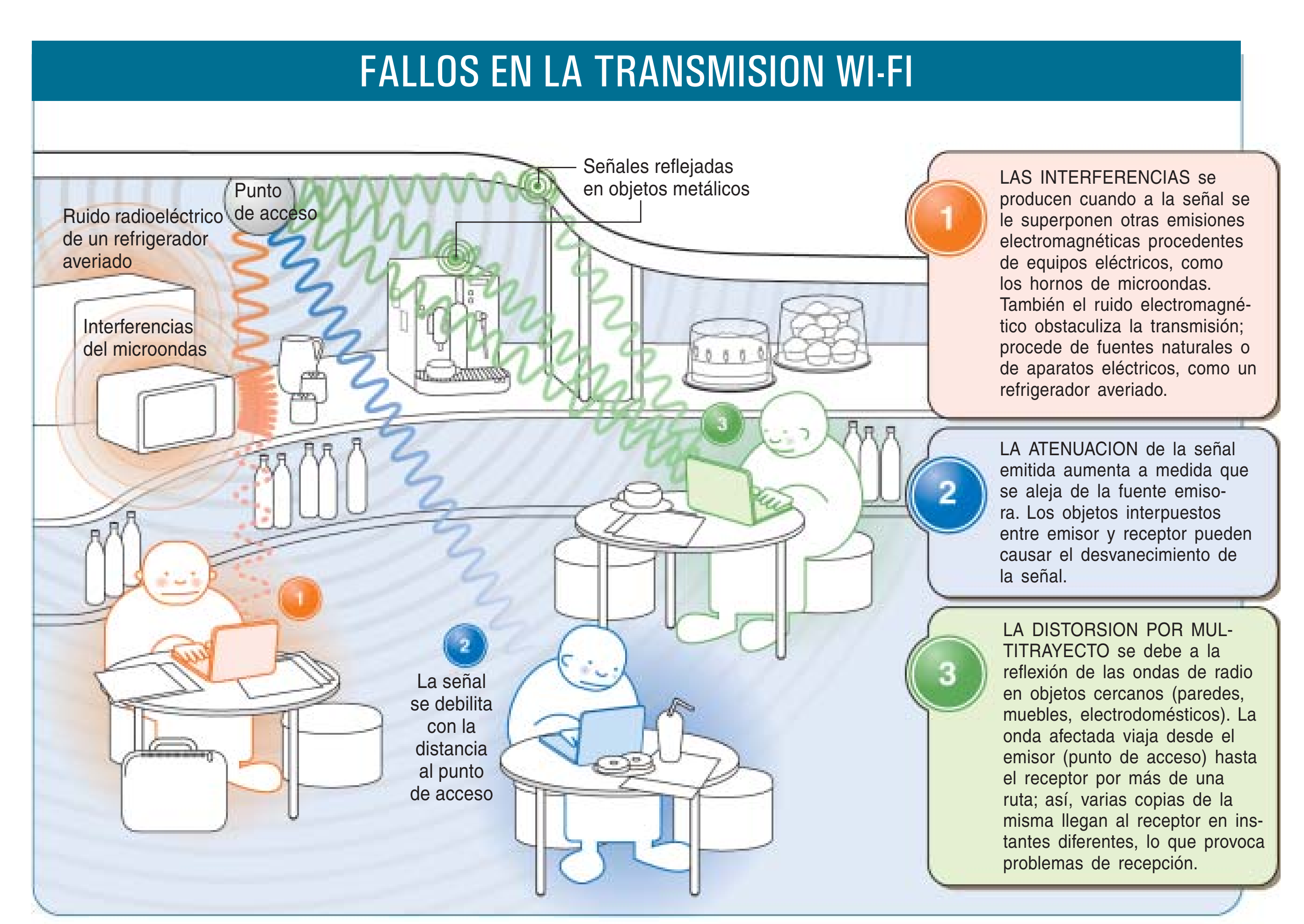
Evitar congestiones

La Wi-Fi adaptativa mejorará el servicio que ofrecen las redes inalámbricas. Aborda, por diversos medios, los problemas de congestión, de alteraciones en el ambiente radioeléctrico y de seguridad.

La congestión de la red se produce cuando coinciden demasiados usuarios en solicitar el servicio de un punto de acceso, provocando la sobrecarga de éste. De ese modo, la congestión se convierte en causa probable de retrasos y degrada el servicio. Dado que el AP de una célula y los clientes que lo utilizan deben compartir un mismo canal de radio (un segmento del espectro electromagnético) y que sólo una estación (un AP o un cliente) transmite con éxito cada vez, se producen colisiones. Las redes Wi-Fi actuales resuelven estos conflictos entre estaciones en competencia mediante el protocolo de acceso múltiple por detección de portadoras con elisión de colisiones (CSMA/CA, de "carrier sense multiple access with collision avoidance").

En el sistema CSMA/CA, antes de enviar una señal, cada estación se halla a la escucha. Si una oye que otra está en el proceso de envío, pospone su acción y espera a que el canal quede libre. Si dos estaciones tratan de enviar una señal casi simultáneamente, ninguna oirá a la otra y sus transmisiones entrarán en colisión. Se impone entonces repetir las transmisiones, ya que ninguna de ellas se recibe correctamente. Cuando muchos ordenadores están usando el mismo punto de acceso,

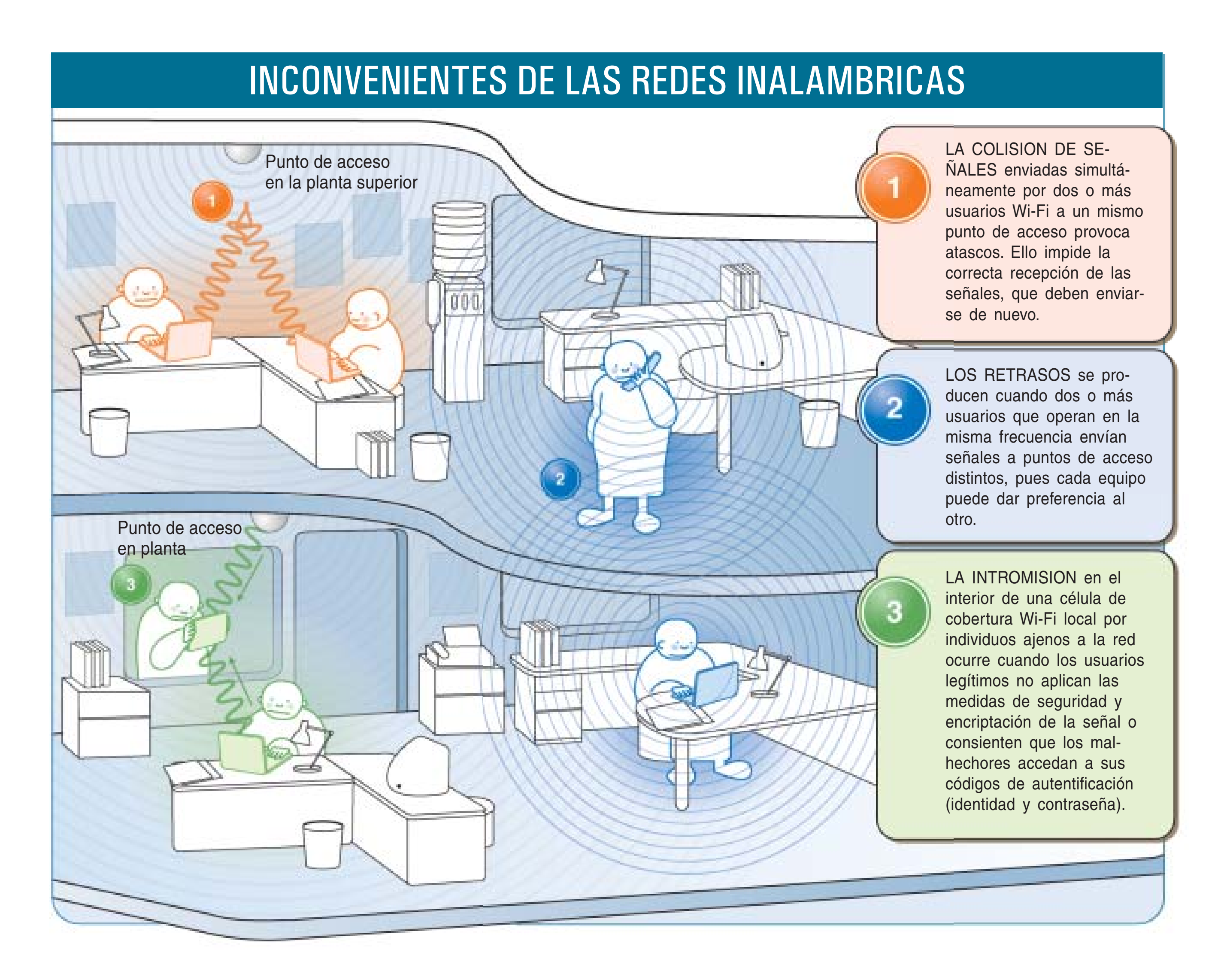
FALLOS EN LA TRANSMISION WI-FI



1. LAS INTERFERENCIAS se producen cuando a la señal se le superponen otras emisiones electromagnéticas procedentes de equipos eléctricos, como los hornos de microondas. También el ruido electromagnético obstaculiza la transmisión; procede de fuentes naturales o de aparatos eléctricos, como un refrigerador averiado.

2. LA ATENUACION de la señal emitida aumenta a medida que se aleja de la fuente emisora. Los objetos interpuestos entre emisor y receptor pueden causar el desvanecimiento de la señal.

3. LA DISTORSION POR MULTITRAYECTO se debe a la reflexión de las ondas de radio en objetos cercanos (paredes, muebles, electrodomésticos). La onda afectada viaja desde el emisor (punto de acceso) hasta el receptor por más de una ruta; así, varias copias de la misma llegan al receptor en instantes diferentes, lo que provoca problemas de recepción.



LUCY READING-IKKANDA

menudean las colisiones y, por ende, habrá que recurrir a múltiples repeticiones de la transmisión; ello acarrea demoras para los usuarios.

El problema de sobrecarga de los puntos de acceso adquiere particular interés en zonas de elevada densidad de usuarios. En Carnegie-Mellon, empezamos a sufrir este problema en las aulas y en las salas de conferencias. Comprendimos enseguida que en estos espacios tan concurridos —con centenares de usuarios de ordenadores móviles— no se podría aspirar, ni de lejos, a las prestaciones de los sistemas cableados.

El protocolo CSMA/CA pierde eficacia también entre puntos de acceso distantes y dispositivos móviles que estén operando en el mismo canal de radio. Pues, si un AP o un dispositivo móvil oye a un cliente o un AP distante (co-canal), se pondrá a la espera, como haría una estación que transmitiera en el interior de su propia célula. Esta superposición co-canal provoca otro tipo de degradación del servicio.

Supongamos, por ejemplo, que Abel y Ana están usando sendos aparatos que operan en el mismo canal de radio, aunque situados en distintos lugares de un edificio y asociados a puntos de acceso diferentes. Si el sistema de Abel oye al de Ana, se pondrá a la espera cada vez que transmita el sistema de Ana; se retrasarán los mensajes que Abel desea enviar. De forma análoga, si el sistema de Ana oye al de Abel, no podrá realizar envíos mientras Abel esté transmitiendo y su servicio de comunicaciones se degradaría. El problema se apreciaría en su crudeza si ambos utilizasen un terminal telefónico.

Estas situaciones se resuelven mediante una cuidadosa asignación de canales y el reparto equilibrado de cargas, una nueva técnica que reduce el riesgo de sobrecarga de un AP. La técnica en cuestión se basa en la posibilidad de que los clientes estén en radio-alcance de dos o más puntos de acceso. Las redes Wi-Fi adaptativas alivian la congestión mediante la distribución equitativa de los clientes entre los AP, para que ninguno de éstos quede sobrecargado, con la mejora consiguiente de la eficacia del sistema.

Cada conexión entre cliente y punto de acceso corresponde a una "asociación". El proceso comienza cuando un cliente inicia una solicitud de asociación. El punto de acceso que recibe la solicitud puede aceptarla o rechazarla. Aunque la norma IEEE 802.11 no especifica ningún algoritmo para la toma de esta decisión, un punto de acceso de segunda generación (o la central de conmutación que lo controla) examina la carga del punto de acceso en cada momento, así como la soportada por los AP próximos, al objeto de facilitar la resolución. Un

AP muy cargado podría no resultar óptimo para la asociación con un cliente nuevo. Si se recibe tal petición y el sistema sabe que en el radio de alcance del solicitante existe otro AP menos cargado, el punto de acceso saturado puede rechazar la solicitud de asociación, reforzando así la eficacia de la red. El reparto equilibrado de cargas, en sinergia con otras técnicas, elevará el rendimiento de las redes Wi-Fi futuras, incluso en lugares donde se dé una concentración nutrida de usuarios.

Entornos radioeléctricos cambiantes

Los problemas de radiodifusión mencionados, a saber, atenuación, recepción multitrayecto, interferencias y ruido, se reducen de forma notable con un proyecto idóneo de la red. Su diseño demanda la resolución de varias cuestiones: cuál es la ubicación de los puntos de acceso (en el interior de un espacio diana) que optimiza la cobertura y el rendimiento, qué canales de radio deben asignarse a cada AP y cuáles son las características del ambiente electromagnético y la geometría del edificio donde se va a instalar la red LAN inalámbrica, al objeto de implantar una red de radiodifusión tridimensional.

La ubicación de los puntos de acceso debe contemplar dos requisitos. Por un lado, conviene no dejar espacios sin cobertura. Por otro, importa que los AP se encuentren tan separados como sea posible para minimizar los costos del equipo y la instalación. La separación de los puntos de acceso reduce también la degradación del rendimiento debida a la superposición de coberturas entre AP que operan en el mismo canal de radio (solapamiento co-canal). En la asignación de canales (la segunda fase del proceso de diseño) lo habitual es minimizar la superposición co-canal, lo que reduce la interacción entre estaciones de células co-canal distintas.

La Wi-Fi adaptativa incorpora otra prestación. Se trata del control automático del tamaño celular: las células se expanden o contraen para adaptarse a las condiciones variables de radiodifusión. Esta técnica permite mitigar los fallos de los AP debidos a diseños poco cuidadosos.

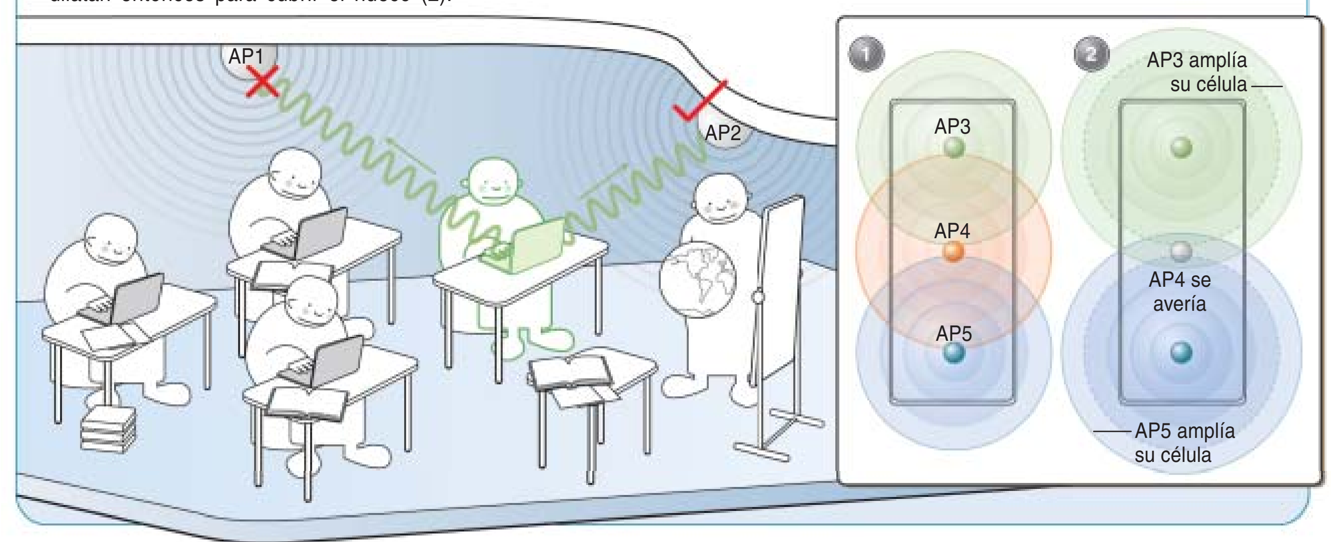
Ni siquiera las redes mejor proyectadas se hallan exentas de posibles alteraciones de las condiciones locales de radiodifusión. Por ejemplo, si en una fábrica se cambian de lugar aparatos metálicos, la modificación de las condiciones electromagnéticas deja huecos sin cobertura. En este caso, resulta conveniente una expansión o contracción adaptativa de las células. El tamaño celular se ajusta mediante la modificación de la potencia de emisión de Wi-Fi en los puntos de acceso. Si las adaptaciones se ajustan con precisión al nuevo entorno de radio, la cobertura continua de la red se mantiene en la totalidad del espacio diana sin superposiciones indebidas de las células. (Los puntos de acceso actuales modifican sólo su propia potencia de emisión, pero las mejoras que ha de recibir el estándar IEEE 802.11 permitirán que los AP instruyan a los clientes para que éstos aumenten o disminuyan la potencia de sus emisiones.)

El control automático del tamaño celular simplifica el diseño de una LAN inalámbrica. No hace falta que la distribución de los AP sea óptima; basta con que sea razonable. Además, y de una forma provisional, sale al paso de los huecos de cobertura que de vez en cuando se originan por culpa de averías en los AP; el sistema reparador actúa en función de la ubicación del punto de acceso y del tipo de antena utilizado.

REDES WI-FI ADAPTATIVAS

Una red Wi-Fi moderna alivia las congestiones mediante la distribución equitativa de las conexiones de usuario entre los puntos de acceso disponibles. Este reparto equilibrado de cargas (*abajo, izquierda*) se activa cuando se intenta establecer conexión con un punto de acceso muy cargado, como el AP1. Si el sistema "sabe" que un segundo punto de acceso, menos solicitado, se encuentra al alcance del usuario, le negará la conexión al primer punto de acceso y le conectará al segundo punto de acceso, lo que aumenta el rendimiento operativo de la red.

Cuando las condiciones de transmisión cambian, un sistema Wi-Fi adaptativo modifica el tamaño de sus células para compensar. En este ejemplo (*abajo, a la derecha*), los AP 3, 4 y 5 proporcionan acceso a un espacio interior (*1*). Si la célula central AP4 falla de forma inesperada, se abre un hiato en la cobertura inalámbrica. Las células vecinas, AP3 y AP5, se dilatan entonces para cubrir el hueco (*2*).



LUCY READING-IKKANDA

Asignación dinámica de canales

En las redes Wi-Fi adaptativas, los puntos de acceso pueden recurrir también a la asignación dinámica de canales para cambiar de canal de forma automática. La asignación de canales intenta minimizar las superposiciones co-canal debidas al entorno de propagación de las ondas de radio. Una vez efectuadas las asignaciones, lo normal es que sean estáticas, por lo que nada garantiza que éstas resulten válidas en el futuro.

Una red Wi-Fi de segunda generación "valora" periódicamente el medio radioeléctrico y luego procede a la reasignación dinámica de canales, según convenga. Ello elimina la necesidad de ejecutar asignaciones de canal durante el proceso de diseño original. Si, por ejemplo, en una oficina se cambia o retira el mobiliario, la zona de cobertura de una célula podría aumentar. Si esta expansión resultase en una superposición de coberturas con otra célula que operase en el mismo canal, se produciría una disminución del rendimiento. En tal caso, quizá convendría conmutar la segunda célula a otro canal. Los algoritmos de conmutación de canal minimizan la superposición de cobertura co-canal en toda la red.

Los sistemas Wi-Fi adaptativos suelen activar de forma periódica un algoritmo de conmutación para asegurar que la asignación de canales refleja la situación radioeléctrica actual. Las técnicas de asignación dinámica mejoran también el rendimiento, al permitir a los puntos de acceso la elección de canales que no sufren los efectos del ruido local o de interferencias.

Seguridad inalámbrica

La inseguridad constituye, con toda probabilidad, el problema estrella de la técnica Wi-Fi. No hay usuario dispuesto a aceptar que extraños puedan hurgar en sus mensajes de correo electrónico o entrar sin permiso en su sistema. El estándar IEEE 802.11 original proporcionaba la encriptación de las transmisiones mediante la opción privacidad equivalente a línea fija (WEP, de "Wired Equivalent Privacy"). La encriptación convierte un flujo de bits en otro flujo (el encriptado), de modo tal que

Comparación entre Wi-Fi y Wi-MAX

Es posible que muchos lectores tengan noticia de WiMAX, una nueva técnica de acceso inalámbrico. ¿En qué se basa? ¿Qué relación tiene con la Wi-Fi?

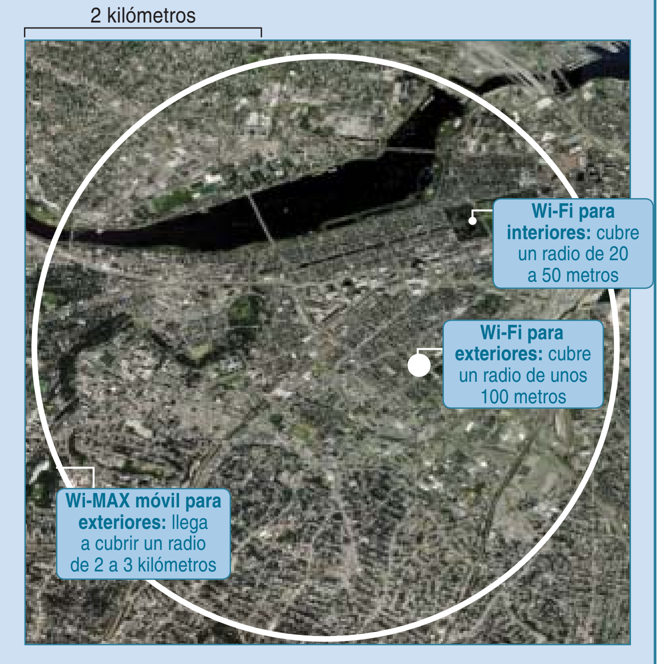
Mientras que Wi-Fi se aplica en dispositivos móviles, WiMAX fue concebida para conexiones a Internet estacionarias. (El término "WiMAX" es creación del grupo industrial WiMAX Forum.)

Wi-Fi se basa en la norma IEEE 802.11. WiMAX proviene de la IEEE 802.16d, que se adoptó en 2004 para definir el servicio inalámbrico de gran velocidad a sedes estacionarias situadas a 50 kilómetros de distancia. Wi-Fi ofrece un alcance máximo de 100 metros. El alcance máximo de WiMAX es claramente superior, porque transmite a potencias mayores, en función de la radiofrecuencia utilizada. WiMAX podrá operar a 75 megabits por segundo (velocidad que supera de lejos la de las líneas ADSL), pero esta capacidad se repartirá entre un gran número de usuarios.

WiMAX se ideó para proporcionar el mismo tipo de servicio rápido de Internet que ofrece el ADSL, el módem-cable y los sistemas de fibra óptica. De ahí su denominación habitual de red de área metropolitana (MAN, de "metropolitan-area-network").

Pese a los orígenes de esta técnica, se ha comenzado a trabajar en una versión móvil de la norma IEEE 802.16e, popularmente conocida por WiMAX móvil. Lo mismo que Wi-Fi, está pensada para dar servicio a ordenadores portátiles y a otros dispositivos móviles. Sin embargo, tendrá mayor alcance; quizá, dos o tres kilómetros.

Pese a los rumores que corren por la industria informática acerca del sistema WiMAX móvil, la norma correspondiente no se ha aprobado todavía. Además, está por ver si esta técnica encuentra asidero en el mercado. Pudiera ocurrir que WiMAX no compitiera directamente con Wi-Fi. Dadas su potencia y alcance mayores, lo más probable es que se conjugue con el servicio celular de tercera generación (3G) para proporcionar servicio móvil de Internet, primero en zonas urbanas y luego en regiones de mayor



extensión. Lo mismo que WiMAX, el sistema de tercera generación opera a mayor potencia y alcance que Wi-Fi.

Es posible que los tres sistemas (Wi-Fi, WiMAX y tercera generación) acaben coexistiendo, ocupando cada uno un "nicho" especializado. Dado que WiMAX y 3G operan a mayores potencias y se valen de métodos de acceso distintos que Wi-Fi, nunca sufrirán los problemas descritos en este artículo.

En el futuro, cada vez será más frecuente que los ordenadores portátiles y los asistentes personales (PDA) estén equipados para operar con diversas redes inalámbricas. Un ordenador portátil podría conectarse mediante Wi-Fi en casa o en la oficina, y utilizar, en cambio, WiMAX o 3G en los demás lugares. Las combinaciones Wi-Fi/WiMAX o Wi-Fi/3G podrían convertirse en procesos rutinarios; seguidos de ordenadores Wi-Fi/WiMAX/3G capaces de establecer conexión con las tres redes.

TERRASERVER.COM/GLOBEXPLORER; LUCY READING-IKKANDA

el original sólo puede recuperarse con la ayuda de una clave, el cifrado utilizado inicialmente para efectuar la codificación. Sin embargo, son muchos los usuarios de sistemas inalámbricos que no se molestan en activar el modo de encriptación, por lo que envían sus mensajes "en abierto", facilitando su interceptación.

Pero la propia WEP tiene sus puntos débiles a través de los cuales pueden descubrirse las claves de la encriptación y, en consecuencia, descifrar los mensajes. La vulnerabilidad de la encriptación WEP pasó a ser de dominio público en 2001. Desde entonces, los expertos se han esforzado en mejorar la seguridad de las redes Wi-Fi.

El acceso autorizado plantea otro problema para las redes Wi-Fi. Los usuarios se identifican ante el sistema mediante un nombre de usuario (ID) y una contraseña. Pero si algún navegante malévolo escucha las transmisiones de otros, fácilmente puede hacerse con la ID y la contraseña de un usuario legal y lograr así acceso a la red.

En 2003 y 2004, el grupo de trabajo de IEEE 802.11 y la Wi-Fi Alliance (el grupo industrial que acuñó el término "Wi-Fi") completaron los trabajos en sendos estándares relacionados entre sí, el IEEE 802.11i y el Wi-Fi de acceso protegido (WPA, de "Wi-Fi Protected Access"), que proporcionan medidas de seguridad más robustas. Incluyen técnicas de encriptación más refinadas y métodos de acceso a las claves de codificación y descodificación (por parte de los AP y los clientes) más seguros.

El sistema WPA (que utiliza otro estándar, el IEEE 802.11X) ofrece también un proceso de certificación más robusto que los disponibles hasta ahora. La combinación de estos nuevos estándares mejora de forma notable la seguridad general de las redes Wi-Fi adaptativas.

Algunos fabricantes de equipos Wi-Fi han añadido otras medidas de seguridad. Pensemos en la detección de intrusos, por ejemplo. Las redes inalámbricas difieren de las redes cableadas en que los dispositivos de escucha subrepticia (e incluso los puntos de acceso) se hallan en cualquier lugar del área de cobertura de una red inalámbrica o cerca de ella. (La intrusión por línea fija puede efectuarse desde gran distancia.) Por este motivo, algunos equipos Wi-Fi se valen de técnicas de localización para detectar la presencia de estaciones ilícitas. El método le permite al sistema encontrar la estación perturbadora y eliminarla.

Con el desarrollo de la Wi-Fi adaptativa, las redes inalámbricas se parecen cada día más a las cableadas. Los usuarios empiezan a percatarse de ello. No obstante, queda todavía mucho por hacer a este respecto. Se trabaja ya en el desarrollo de un método de localización automática de dispositivos móviles en una red Wi-Fi. Esta técnica resultaría de gran utilidad cuando fuera necesario localizar con prontitud personas (médicos en un hospital), así como objetos (productos que se desplazan por una línea de montaje en una fábrica).

El desarrollo y la evolución de Wi-Fi y otras técnicas de comunicación inalámbrica se están produciendo a un ritmo vertiginoso. Cada vez son más quienes abandonan los servicios de telefonía fija y optan por los teléfonos celulares móviles. Algunos gobiernos municipales, como el de Filadelfia, están creando amplias zonas con cobertura Wi-Fi. Al propio tiempo, ha despegado la telefonía celular de tercera generación; Wi-Max, una nueva técnica inalámbrica, podría muy pronto hacerse con un importante segmento del mercado. Vivimos en un mundo cada vez más inalámbrico.

Fabricantes de equipos Wi-Fi adaptativos

COMPAÑIA	UBICACION	SITIO EN LA RED
Aruba Networks	Sunnyvale, California	www.arubanetworks.com
Cisco Systems/ Airespace*	San José, California	www.airespace.com
Cisco Systems/ Aironet*	San José, California	www.cisco.com
Colubris Networks	Waltham, Massachusetts	www.colubris.com
Extreme Networks	Santa Clara, California	www.extremenetworks.com
Symbol Technologies	Holtsville, Nueva York	www.symbol.com
Trapeze Networks	Pleasanton, California	www.trapezenetworks.com

**Cisco Networks ha adquirido Airspace, una compañía de equipos Wi-Fi adaptativos. Aironet, el producto Wi-Fi de Cisco, incorpora ya prestaciones adaptativas.*

El autor

Alex Hills es catedrático de ingeniería eléctrica e informática en la Universidad Carnegie-Mellon. Su labor docente e investigadora se centra en las técnicas inalámbricas y en las políticas de telecomunicación.

Bibliografía complementaria

WIRELESS ANDREW. Alex Hills en *IEEE Spectrum*, vol. 36, n.º 6, págs. 49-53; junio de 1999.

LARGE-SCALE WIRELESS LAN DESIGN. Alex Hills en *IEEE Communications*, vol. 39, n.º 11, págs. 98-107; noviembre de 2001.

REAL 802.11 SECURITY: WI-FI PROTECTED ACCESS AND 802.11I. John Edney y William Arbaugh. Addison-Wesley Professional, 2003.

ROLLABOUT: A WIRELESS DESIGN TOOL. Alex Hills y Jon Schlegel, en *IEEE Communications*, vol. 42, n.º 2, págs. 132-138; febrero de 2004.

RADIO RESOURCE MANAGEMENT IN WIRELESS LANS. Alex Hills y Bob Friday en *IEEE Communications*, vol. 42, n.º 12, págs. S9-S14; diciembre de 2004.

WIRELESS NETWORKS FIRST-STEP. Jim Geier. Cisco Press, 2004.

THE IEEE 802.11 HANDBOOK: A DESIGNER'S COMPANION. Segunda edición. Bob O'Hara y Al Petrick. IEEE Press, 2005.

802.11 WIRELESS NETWORKS: THE DEFINITIVE GUIDE. Segunda edición. Matthew Gast. O'Reilly, 2005.

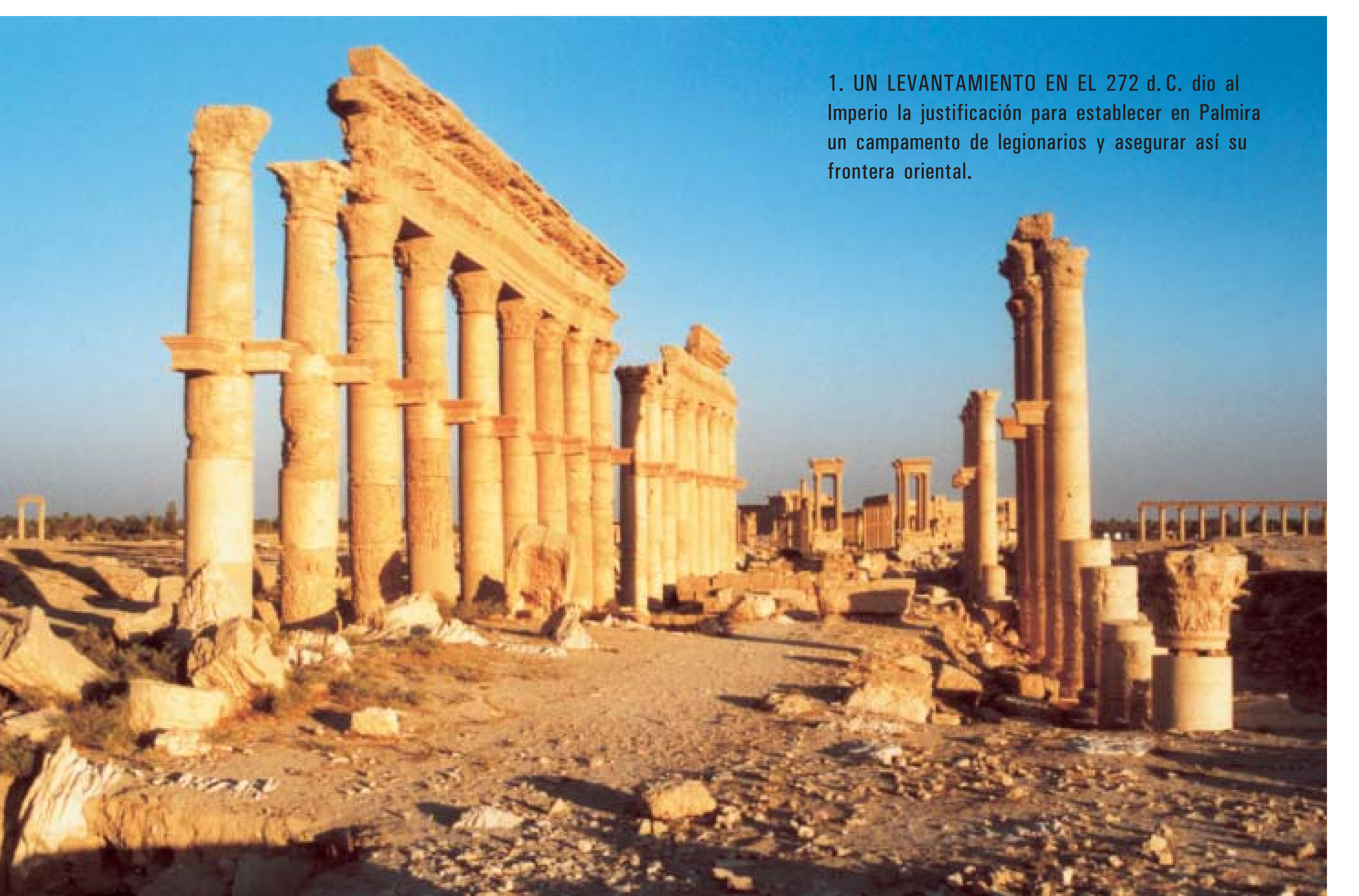
La guerra del desierto

Casi tan pronto como acabó su conquista, quedó claro que las provincias romanas de Oriente iban a ser difíciles de mantener. Varios emperadores romanos fracasaron en su lucha contra unos pueblos persas levantiscos

Theodor Kissel

Palmira, Dura Europos, Ctesifonte... nombres sonoros de los extremos orientales del mundo romano. Hasta la actual Siria llevó el Imperio sus legiones. Sin embargo, la situación fue estable en aquellas lejanas fronteras sólo durante un corto tiempo. Surgieron peligrosos enemigos, en especial en lo que ahora es Irán: los partos primero y luego los sasánidas. Los arqueólogos han encontrado en estos últimos años huellas de derrotas devastadoras: torres defensivas derruidas, osamentas de legionarios romanos en túneles defensivos o vestigios de incendios. Pero también han hallado medios de defensa más exitosos: castillos en posiciones adelantadas que podían ser defendidos con mayor facilidad, o campamentos situados tras el *limes* oriental (la frontera fortificada en ese confín del Imperio), desde donde podían desplazarse rápidamente unidades móviles a los puntos conflictivos.

La implicación de Roma en Oriente comenzó muy poco después del término de la Segunda Guerra Púnica (218-201 a. C.). En el ínterin, la metrópoli tiberina había dominado la península itálica por completo, vencido a la odiada Cartago y emprendido el sometimiento de España. Pero también aparecieron por entonces otras grandes potencias en el Mediterráneo oriental. Protagonizó el primer estallido el joven macedonio Filipo V, quien aspiraba al predominio en el ámbito cultural griego.



1. UN LEVANTAMIENTO EN EL 272 d. C. dio al Imperio la justificación para establecer en Palmira un campamento de legionarios y asegurar así su frontera oriental.

2. COMPACTO Y FACIL DE DEFENDER era el nuevo tipo de castillo de la Antigüedad tardía. La imagen muestra el de Qasr Bushir, en la Jordania actual. Se construyó en el tiempo del emperador Diocleciano.



Roma mandó sus legiones y sofocó esas ambiciones de raíz en el año 197 a. C. Sin embargo, no habían transcurrido cinco años cuando el monarca seléucida Antíoco III (*véase el recuadro* "El imperio de los aqueménidas y los seléucidas") extendió sus tentáculos hasta el Asia Menor e incluso Grecia. Sobrevaloró sus fuerzas y finalmente los seléucidas tuvieron que conformarse con unos restos miserables, que acabaron también siendo anexionados al Imperio, en el 64 a. C. Roma llegó entonces al Eufrates, el río que durante varios siglos iba a señalar su frontera oriental (*véase la figura 3*).

Es cierto que no faltaron intentos de continuar expandiéndose. Pero al otro lado del río se encontraban los dominios de los partos (*véase el recuadro* "El imperio de los aqueménidas y los seléucidas"), que no sólo infligieron una dolorosa derrota al general romano Craso en el 53 a. C., sino que cruzaban el Eufrates con cierta molesta asiduidad. Estos nuevos enemigos resultaban mucho más insidiosos; el fortalecimiento de la frontera era una tarea más urgente que la consecución de nuevos territorios. Augusto (63 a. C.-14 d. C.), el primer emperador romano, envió las legiones romanas a Oriente. Bajo los emperadores Flavios (que gobernaron del 69 al 96 d. C.) se erigieron campamentos de legionarios, castillos y puestos de observación en las orillas del Eufrates, unidos mediante una vía fortificada. Este *limes* se extendía desde Trapezus, a orillas del Mar Negro, hasta Sura, en el norte de Siria, y era completamente distinto del *limes* germánico superior-rético. No estaba completamente fortificado a lo largo de toda su extensión; las defensas se concentraban en sitios estratégicos, como los vados y los pasos de los ríos. Uno de esos lugares era Zeugma, hoy famoso por las excavaciones internacionales con que la Unesco ha querido poner a salvo sus restos, condenados por la construcción de un pantano.

La relación entre romanos y partos siguió siendo bastante tensa durante el siguiente siglo. Tan pronto primaba la coexistencia amistosa como buscaba una de las partes la supremacía militar; esto último ocurrió con los emperadores Trajano (que gobernó del 98 al 117 d. C.) y Lucio Vero (del 161 al 166 d. C.). Sin embargo, nada conseguiría variar el equilibrio de poder en Oriente en aquellos tiempos.

Tras las huellas de Alejandro

La situación sólo cambió a comienzos del siglo III d. C., cuando a los emperadores de la dinastía de los Severos (que gobernaron del 192 al 235 d. C.) les obsesionó una idea fija: la *imitatio Alexandri*, la imitación de Alejandro, como rezaba su lema. Siguiendo las huellas de Alejandro, los generales romanos aprovecharon las debilidades del enemigo, en especial las luchas intestinas que se desencadenaron en aquellos días por la posesión del trono, y avanzaron profundamente hacia el este, hasta llegar ante las mismas puertas de Ctesifonte, la capital de los partos. Los funcionarios de la administración pudieron establecer así dos nuevas provincias: Osrhoena y Mesopotamia. Se pertrecharon tres nuevas legiones y se establecieron campamentos fortificados también al otro lado del Eufrates. El *limes* romano de Oriente corría ahora a lo largo de la línea que unía el río Chabur con Gebel Singar y el Tigris.

Esta fase de la política oriental romana se había conocido hasta ahora bastante mal. Hace pocos años, investigadores alemanes localizaron varios asentamientos militares mediante fotografías aéreas y empezaron a excavarlos. Uno de ellos era Markus Gschwind, del Instituto Arqueológico Alemán con sede en Damasco. Puso al descubierto el campamento

romano de Qreiye, a doce kilómetros de la moderna ciudad siria de Deir ez-Zor. Con la ayuda de la prospección geomagnética y geoeléctrica, los arqueólogos localizaron las murallas del castillo enterradas y excavaron de forma ordenada y provechosa. El edificio cuadrangular, levantado sobre una colina de basalto que domina el Eufrates, medía 220 metros de lado. Probablemente fue mandado construir por el emperador Septimio Severo para la campaña parta (195-198 d. C.). Todo parece cuadrar con esta hipótesis: la fortaleza, según los restos de cerámica hallados, se erigió poco antes de 200 d. C. Con su sistema de doble muro con foso interpuesto, su muralla defensiva con cuatro puertas, cuatro torres esquineras y otras ocho más situadas entre ellas, como muestran los restos, coincide con el patrón dominante en los dos primeros siglos de la era cristiana. Desde hace poco, se conocen los nombres de 1500 soldados que estuvieron acantonados en esa plaza fuerte: se ha encontrado en el edificio del estado mayor (*principia*) un bloque de piedra yesosa de alrededor de dos quintales de peso con una inscripción que informa sobre la ceremonia de consagración de una tropa auxiliar (*auxilia*) procedente de Petra, la ciudad de piedra en la actual Jordania.

Petra, capital de los nabateos, se encuentra mucho más al sur, a más de 1000 kilómetros de distancia, pero era costumbre de la administración militar romana desplazar tropas auxiliares a los terrenos sometidos y estacionarlas tan lejos de su patria como fuera posible. Una estrategia simple, pero efectiva. Pues sin el apoyo de su propio pueblo, los soldados guardaban lealtad al Imperio, que les garantizaba la manutención. Sin embargo, con frecuencia estas tropas vivían, por así decirlo, en ninguna parte. Así sucedía en Qreiye: la ciudad más próxima, Palmira, se hallaba a ocho o diez días de viaje.

El imperio de los aqueménidas y los seléucidas

La estirpe de los sasánidas fue la última en la sucesión de dinastías soberanas "persas" que desde mediados del último milenio antes de Cristo hasta la conquista islámica del año 642 se establecieron en suelo del actual Irán. Desde allí unificaron en un imperio a los distintos pueblos asiáticos limítrofes.

Los primeros en dicha sucesión fueron los aqueménidas (550-330 a. C.), cuyo gigantesco imperio se extendió desde el Indo hasta el mar Mediterráneo y desde los desiertos libios hasta las estepas de Asia central. Los soberanos aqueménidas, originarios de la provincia central de Irán, Persis, se consideraban los elegidos por el dios Ahuramazda y sus representantes en la tierra. Desde Persépolis, "la ciudad más rica bajo el sol", gobernaba el "rey de reyes" su imperio. A pesar de su poderío, Alejandro Magno, que irrumpió en Asia en el 334 a. C., conquistó Persia en unos pocos años.

Tras la temprana muerte de Alejandro, sus generales se repartieron su imperio y uno de ellos, Seleuco I, fue el que asumió el poder en Asia. Fundó la dinastía de los seléucidas (312-64 a. C.). Alrededor del 250 a. C. irrumpieron en el imperio los parnos, un pueblo seminómada procedente de las estepas al sur del mar de Aral, y ocuparon la provincia de Partia, de ahí su posterior nombre de partos.

Desde principios del siglo II a. C., los seléucidas veían inevitable la confrontación en Occidente con la emergente potencia romana. No contrarrestaron en una medida suficiente los ímpetus expansivos de los partos. A consecuencia de este abandono, cayeron en manos de los partos Seleucia, en el 141 a. C., y por fin, en el 129 a. C., toda Mesopotamia. Así extendieron su poder al otro lado del Eufrates. Cuando el Imperio Romano destronó al último monarca seléucida, se desató la guerra entre romanos y partos.



El imperio de los seléucidas y sus vecinos alrededor del 230 a. C.

EMDE-GRAFIK/SPEKTRUM DER WISSENSCHAFT, SEGUN THEODOR KISSEL

Dura instrucción en el frente

Los traslados a campamentos vecinos o los viajes de servicio al palacio del gobernador suponían, en esas circunstancias, una auténtica liberación de la rutina. Y si se tenía suerte, hasta podían conseguirse unos días de vacaciones —*commeatus* en latín—, bien para poner en orden cualquier asunto personal, bien por una visita de la familia (sobre esto encontramos información en un papiro procedente de Karanis, localidad del Bajo Egipto). O por convalecencia: el servicio en los campamentos apartados de la civilización era durísimo. Del rigor imperante en aquellos campamentos nos informan los planes de servicio de la vigésima cohorte de Palmira, encontrados en Dura Europos. Trabajos de zapa en las trincheras agotadores, instrucción militar y guardias continuas a temperaturas que superaban los cuarenta grados a la sombra, siempre con la vista puesta en el enemigo que acechaba al otro lado del Eufrates.

También la consecución de los suministros ocasionaba costosos trabajos. Con el sudor de su frente, nunca mejor dicho, los soldados cultivaban cereales y verduras en los fértiles

terrenos de aluvión del Eufrates, compraban o intercambiaban alimentos con los nativos y requisaban en los lugares de los alrededores legumbres o se las compraban a los comerciantes que llegaran hasta aquellos apartados lugares.

Apenas treinta años de media alcanzaban los soldados romanos destacados en el frente, si nos atenemos a las inscripciones funerarias repartidas por todo lo largo y ancho del Imperio Romano. En los páramos de Qreiye sólo podían sobrevivir los hijos del desierto, como los nabateos, que además tenían ya experiencia sobrada en la guerra en un medio tan hostil.

Las siguientes huellas de la presencia romana más allá del Eufrates se encuentran 120 kilómetros hacia el nordeste. Provienen de un antiguo asentamiento asirio en la colina de Dur-Katlimmu (Seh el-Hamad en la actualidad). Allí, a orillas del Chabur, Hartmut Kühne encontró, sepultados bajo unos cuantos metros de tierra, los restos de un campamento romano; las monedas desenterradas muestran que data de la misma época que Qreiye. Es muy probable, además, que se trate de Magdala. Este lugar sólo se conocía por textos conservados en los papiros de Dura Europos, donde se reclutaba la guarnición de Magdala. Y con este campamento toma forma, a comienzos del siglo III d. C., la línea defensiva romana en Oriente, a la que también pertenecían las plazas fuertes de Zagurae (Ain Sinu), situada al este de Gebel Sinjar, y Castra Maurorum (Seh Qubba), 100 kilómetros al nordeste de Mosul.

La presencia militar se hacía allí apremiante. En el año 224 d. C., en antiguo territorio persa, en la región situada en torno a la actual Shiraz, ciudad del sudoeste de Irán, un vasallo de los reyes partos, Ardashir, se había levantado en armas contra sus señores. Desde las expediciones de conquista llevadas a cabo por Alejandro Magno, los pueblos de las tierras sometidas habían observado pasivamente la expansión del espíritu griego. Impulsado por un movimiento nacionalista persa, Ardashir fundó la dinastía de los sasánidas (nombre que deriva de Sasán, sacerdote zoroástrico y abuelo de Ardashir, que ejercía su ministerio cerca de Persépolis a comienzos del siglo III a. C.). Ardashir alardeaba de pertenecer a



3. PARA ASEGURAR LA REGION CRITICA del Eufrates, el Imperio estableció en dos fases su *limes* (frontera fortificada) oriental.

la tradición de la dinastía de los aqueménidas, cuyos reyes, partiendo de Persia, fundaron cientos de años antes un imperio que se extendió por medio continente (*véase el recuadro* "El imperio de los aqueménidas y los seléucidas").

"Lo que vosotros, romanos, ocupáis en Asia, es mi herencia"

La religión y los valores iranios sustituyeron a los ideales helenistas que habían dominado con los seléucidas y los partos. Los reyes sasánidas, tal y como habían hecho sus predecesores aqueménidas, se rodearon de un aura de divinidad. Como representantes en la tierra de Ahuramazda, el dios supremo del panteón iranio, se creían provistos del *farnah*, el "resplandor divino" otorgado por este dios, que los situaba muy por encima de los demás nobles.

Esta elección por Ahuramazda tuvo expresión gráfica en los llamados "relieves de investidura". Orgullosamente sentados en un corcel, recibían los sasánidas de las manos de Ahuramazda el anillo que les otorgaba el poder soberano. Para resaltar todavía más su legitimación divina y su poder absoluto, mandaron esculpir estos relieves en Naqsh-i Rustam, el lugar donde se enterraba a los reyes aqueménidas. Ante los monumentos de sus supuestos ancestros, pregonaban en los relieves su autoproclamación como "reyes de reyes", título que ya habían llevado sus lejanos predecesores, los soberanos aqueménidas.

Siguiendo esas tradiciones, los sasánidas instauraron un imperio estructurado feudalmente, con un fuerte poder central. Sin embargo, el reconocimiento de la herencia de los antiguos persas no se limitó a la política interior. "Lo que vosotros, romanos, ocupáis en Asia, es mi herencia", manifestó Ardashir apenas transcurridos seis años de su acceso al poder. Con estas palabras dio comienzo una contienda entre las dos grandes potencias que iba a durar más de 400 años.

Ante la llegada de estas noticias intranquilizadoras, las guarniciones de Qreiye y Magdala, así como los restantes puestos fronterizos del

EMDE-GRAFIK/SPEKTRUM DER WISSENSCHAFT, SEGUN THEODOR KISSEL

Eufrates y el Chabur, debieron de ponerse en estado de alerta. Cuando la ciudad de caravanas de Hatra, ante la amenaza sasánida, decidió pedir la ayuda de Roma alrededor de la mitad de la década de los años treinta del siglo III, la superpotencia occidental envió a Hatra la cohorte *XI Maurorum Gordiana*, una unidad formada por norteafricanos, habituada al desierto. A pesar de las medidas de precaución que tomaron, las legiones romanas no tardaron en verse forzadas a ponerse a la defensiva contra los ataques de los jinetes persas, primero en Mesopotamia y luego en Siria. La acción táctica conjunta de la caballería, fuertemente armada con la caballería ligera de los arqueros puso en apuros a los romanos una y otra vez.

Los legionarios habían sufrido numerosas bajas en la batalla de Carrhae contra los partos. "Siempre que aparecen los arqueros, sus flechas perforan los blancos elegidos", dejó escrito el historiador Plutarco (50-120 d. C.), rememorando aquel episodio. Como miles de veces antes, el Imperio confió en formaciones defensivas, por ejemplo la "tortuga" (*testudo*), en la cual los legionarios se colocaban agachados muy juntos unos de otros, protegiendo sus filas mediante escudos situados sobre sus cabezas y a los lados de la formación. Armaban así una suerte de caparazón que los cubría por completo. Sin embargo, a pesar de todos sus esfuerzos, las flechas persas perforaban sin mayores problemas ese techo de hierro.

La fuerza de penetración de aquellas flechas se determinó ya explícitamente en la Antigüedad. Se expone un procedimiento para medir el impulso de una flecha en el manual *De arcu* (Sobre el arco), dedicado, como indica su nombre, a servir de introducción al tiro con esa arma. Se redactó en el siglo II d. C., pero sólo se conserva fragmentariamente en otro escrito de alrededor del 600 d. C., *De re strategica* (Sobre la estrategia), obra del *magister Syrianus*, teórico militar bizantino. Por desgracia, el pasaje recogido se interrumpe antes de que se describa el modo de efectuar el cálculo.

En la corte del rey Abgar VIII de Osrhoene, que fue temporalmente provincia romana, parece que se midió de forma experimental, en torno al año 210 d. C., la velocidad alcanzada por una flecha; y por conocedores en la materia: un escita y un parto. Según el escritor cristiano Sexto Julio Africano (160/170-240 d. C.), diez arqueros pertrechados con arcos persas habrían disparado en aquella ocasión repetidas veces sus flechas, de tal forma que cada uno disparaba su proyectil en cuanto la flecha del precedente se había clavado en el blanco. Cada flecha recorría un sexto de estadio, es decir, un sexto de 185 metros. En total, las flechas disparadas en una hora recorrieron una distancia de 1000 estadios. Julio Africano añade a esos 24.000 estadios por día 1000 estadios más para compensar el tiempo de reacción de los arqueros. Traduciendo a las medidas actuales estos cálculos, se tiene que las flechas persas alcanzaban una velocidad de 192,7 kilómetros por hora.

Semejante velocidad se explica por la forma de construcción de los arcos persas, sumamente flexibles: varias capas de distintos materiales —madera, cuerno, tendones animales—, adheridas unas a otras en un trabajoso proceso, otorgaban al arco una elevada elasticidad. La ventaja de los tendones y el cuerno residía en su capacidad de almacenar energía y transmitirla a continuación a la flecha. A esto se añadía que el arco flexible estaba tensado en sentido opuesto a su curvatura natural, con lo que se multiplicaba su capacidad de penetración.

Acorazados con cuarenta kilogramos de bronce

No menos temibles eran los caballeros acorazados (*clibanarii*) persas, fortalezas móviles comparadas por los autores antiguos con estatuas de bronce ecuestres. Embutidos en un blindaje de plaquitas de bronce unidas entre sí mediante alambres, irrumpían a lomos de sus corceles en las líneas enemigas dejando tras de sí un rastro de pavor allí por donde pasaban.

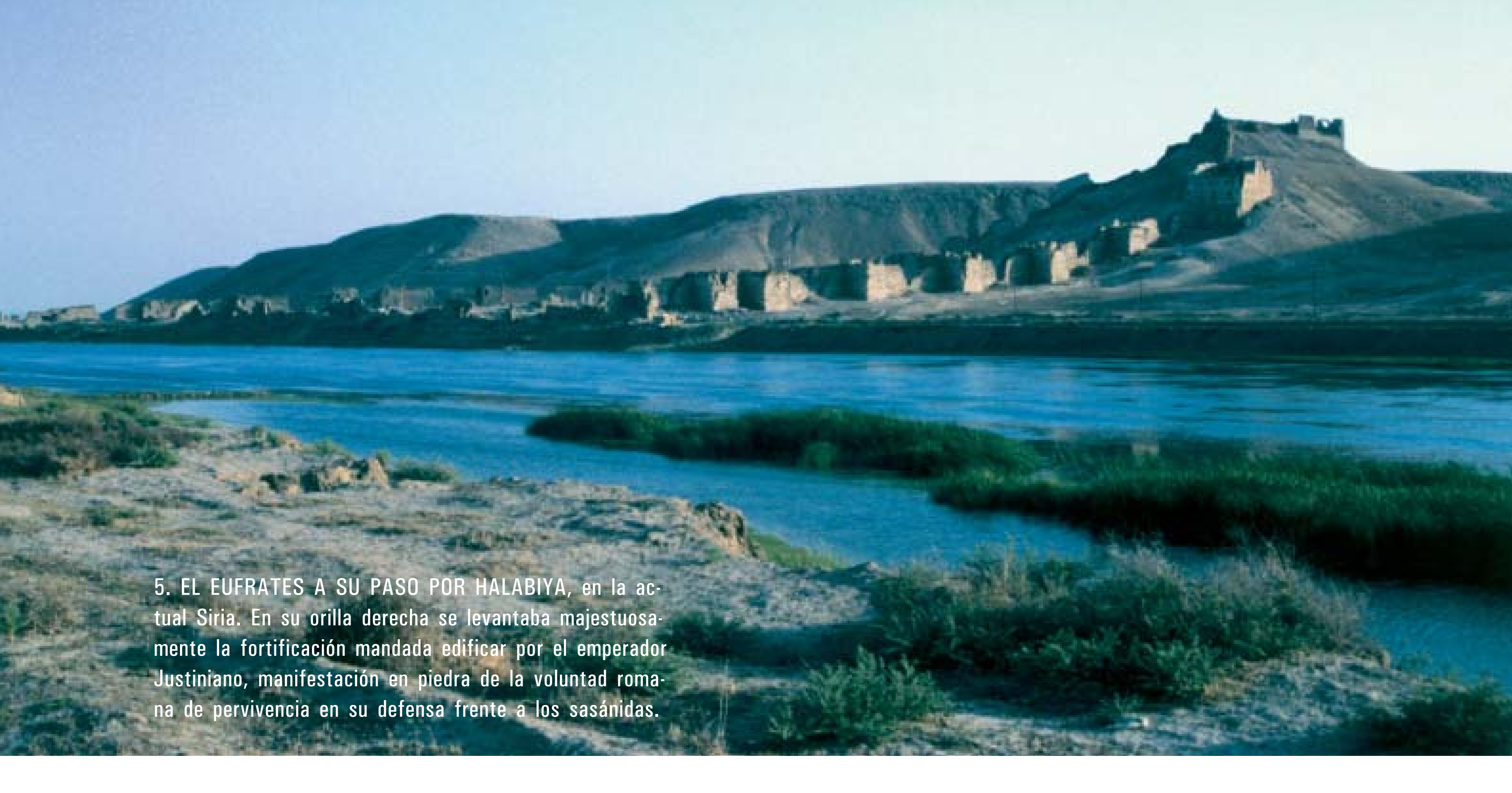
Heliodoro (siglo III d. C.), originario de Siria, describió el blindaje y los efectos de los *clibanarii* en su novela *Aitiopika*, (Etiópicas [IX, 15]): "La armadura se consigue de la siguiente manera: se ensamblan, unas sobre otras, placas forjadas cuadrangulares de aproximadamente el tamaño de un palmo, de tal manera que queden imbricadas tanto horizontal como verticalmente y, a continuación, se cosen por sus bordes. Se fabrica así una envoltura escamada que se amolda cómodamente al cuerpo y rodea completamente los miembros, protegiéndolos sin impedir los movimientos (...). Acorazado sube entonces el jinete al caballo. Para subirse a la montura, debe ser ayudado. Al entrar en combate, deja sueltas las riendas, pica las espuelas y se abalanza sobre el enemigo con un estrépito propio de un hombre de hierro o una estatua metálica viviente. La lanza, portada horizontalmente y sobresaliendo hacia delante con su punta bien visible, se asegura mediante un lazo al cuello del caballo. El extremo del asta está sujeto mediante un cabestrillo colocado en las ancas del caballo, impidiendo que la lanza ceda con el choque y protegiendo además la mano del jinete, quien sólo ha de guiar el golpe. Cuando el jinete enristra la lanza y embiste con violencia, atraviesa cuanto se interpone en su camino y traspasa a veces a dos hombres con un solo golpe." Esta arma de guerra —precedente de los caballeros del occidente medieval— sólo era eficaz en un campo de batalla abierto, dada su limitada movilidad (según apreciaciones actuales, el equipamiento de un *clibanarius*, incluidas las armas, pesaba unos cuarenta kilogramos).

El Imperio, cada vez más debilitado a causa de las guerras internas, no contaba con nada capaz de contrarrestar la mortífera combinación de la caballería acorazada y los arqueros a caballo. Desde que el último soberano de la dinastía de los Severos cayó en el 235 d. C. asesinado por sus soldados, ningún emperador se mantuvo en el poder más allá de unos pocos años (en el 238 d. C. hubo seis emperadores distintos en el trono). Todo ello hizo que el sistema defensivo romano se derrumbara poco después como si fuera un castillo de naipes. Alrededor del 240 d. C. cayeron Hatra y Nisiblis, capital ésta de la provincia de Mesopotamia, que había sido establecida como tal unos cuantos años antes. Tras una contraofensiva emprendida bajo el emperador Gordiano III (242-244 d. C.), que condujo las legiones siguiendo



4. LA MURALLA SUR DE DURA EUROPOS era un punto débil de la fortificación. Los sasánidas consiguieron derruirla parcialmente mediante la excavación de un túnel.

THEODOR KISSEL



5. EL EUFRATES A SU PASO POR HALABIYA, en la actual Siria. En su orilla derecha se levantaba majestuosamente la fortificación mandada edificar por el emperador Justiniano, manifestación en piedra de la voluntad romana de pervivencia en su defensa frente a los sasánidas.

el curso del Eufrates hasta Babilonia, Shapur I, hijo de Ardashir, derrotó a los invasores, al menos según las inscripciones persas, el 244 d. C. en Misiche (Faluya). El emperador murió allí —de nuevo según los persas— y su sucesor Filipo I (244-249 d. C.) compró con dinero una ignominiosa paz y retiró sus legiones a Siria.

En el año 256 d. C. los sasánidas conquistaron Dura Europos, el bastión romano más importante a orillas del Eufrates. Aquí los persas demostraron que se habían convertido en maestros en el arte del asedio. La poderosa fortificación tenía un punto débil muy evidente: su flanco sur estaba abierto al desierto. En esa zona, el terreno arenoso permitía socavar la fortificación.

Lucha a vida o muerte bajo tierra

Todavía hoy se pueden reconocer los túneles de asedio (*cuniculi*, conejeras) con que los persas consiguieron abatir torres y parte de los muros de la fortificación. Este método de asedio fue descrito por el escritor militar romano Vegencio, en torno al 400 d. C., en su *Del arte militar*, con las siguientes palabras: “Cuando los atacantes han conseguido alcanzar los fundamentos de los muros de la ciudad, socavan allí un tramo lo más grande posible y colocan debajo soportes de madera seca. Con esta construcción provisional retardan la caída de los muros. Traen entonces leña y otros materiales combustibles y los colocan también allí. Finalmente disponen sus tropas para el ataque y prenden fuego al andamiaje. Tan pronto como los soportes y el encofrado se queman, el muro se derrumba de repente y queda franca una vía para el asalto”.

Como las excavaciones en Dura Europos han puesto de manifiesto, los romanos no permanecieron pasivos ante esa labor de zapa. Excavaron un contratúnel bajo una de las torres de defensa. Esta forma de contraataque había sido propuesta por Aineas Taktikos, experto griego en asedios en el siglo IV a. C. Sin embargo, por desgracia para la guarnición, esta vez la empresa no se vio coronada con el éxito. Los esqueletos y utensilios encontrados en el túnel no dejan lugar a dudas sobre lo sucedido. Resulta evidente que hubo una cruenta lucha bajo tierra y que los romanos huyeron precipitadamente, sin poner siquiera a salvo a los heridos. Cegaron a toda prisa su propio túnel. Markus Gschwind describió así aquel último golpe contra Dura Europos: “Cuando el fuego de los sasánidas alcanzó las vigas de madera, el fundamento cedió. Dos caras de la torre y un tramo del muro de defensa se hundieron de golpe...”

A Dura Europos le siguieron los campamentos de Qreiye y Magdala, que poco tiempo después fueron conquistados por los guerreros de Shapur. Es muy probable que Qreiye sea Birta Aupan, cuya conquista se recoge en la crónica de los hechos de Shapur, el *Res gestae Divi Saporis*. Pero lo que estaba por venir fue todavía peor. En el 260 d. C. el emperador Valeriano (253-260 d. C.) fue hecho prisionero y encarcelado; murió en prisión, ignominia sin parangón en la historia del Imperio. Shapur I hizo inmortalizar su victoria sobre los romanos en relieves triunfales de tamaño descomunal, distribuidos por su imperio. Bien visibles para todas las miradas, mayestáticos, representan la humillación de tres emperadores romanos: Gordiano III, que yace a los pies del corcel de Shapur, Filipo I (el Arabe), arrodillado, y Valeriano “agarrado” por las manos como un súbdito.

La imagen de una Roma capaz de poner orden en Oriente quedó maltrecha. La ciudad siria de Palmira, bajo el gobierno de su reina Zenobia, se rebeló contra la Urbe en el 272 d. C. Pero Palmira y su reina eligieron un mal momento para hacerlo: en el mismo año falleció el rey sasánida y Roma recibió un respiro inesperado que supo aprovechar. Su nuevo emperador Aureliano (270-275 d. C.) acabó con la escapada de Zenobia y envió una legión a Palmira; sería un primer paso para la estabilización de la situación en la frontera oriental.

THEODOR KISSEL

Un nuevo orden global, sin embargo, sólo lo emprendería un poco más adelante el emperador Diocleciano (284-305 d. C.), quien, gracias a una política inteligente, consiguió por primera vez en mucho tiempo que volvieran a reinar la paz y el orden en el Imperio. Reforzó las defensas fronterizas orientales sobre una nueva base: fábricas de armas estatales situadas en los territorios interiores, y por lo tanto a buen resguardo de los ataques enemigos, proveían a un sistema escalonado formado por puestos de control avanzados y castillos situados en la retaguardia. Diocleciano no quería una frontera herméticamente cerrada, sino un sistema que permitiera detectar enseguida a los invasores y enviar tropas a los lugares críticos. La columna vertebral estratégica del nuevo *limes* la constituyó la *Strata diocletiana*, una calzada militar que se extendía desde Sura, a la vera del Eufrates, hasta Damasco.

Burgos compactos, tropas móviles

Michaela Konrad ha excavado varios castillos a lo largo del tramo norte de la *Strata diocletiana.* Las edificaciones se corresponden con un nuevo tipo de castillo, el *quadriburgium* (*véase la figura 2*), que en Oriente empezó a entrar en uso a finales del siglo III d. C. Este fuerte de la Antigüedad tardía tenía una base trapezoidal y unos bastiones esquineros que sobresalían marcadamente del resto del perímetro, casamatas a salvo de los disparos de flechas, muros dotados de adarves y una ancha muralla defensiva que lo circunvalaba. Los bastiones de las esquinas, de adobe incombustible, dejaban al enemigo poca superficie de ataque y la guarnición no estaba ya en la línea de tiro de las flechas de los arqueros persas. Además, proporcionaban también un modo efectivo de combatir la excavación de túneles, pues desde ellos podían dispararse flechas a quienes los estuviesen abriendo, bien cuando entrasen o bien cuando saliesen de la galería.

Al hacer más pequeñas las edificaciones defensivas, se redujo la cantidad de las tropas a su servicio. Las legiones pasaron ahora a componerse de sólo 1000 hombres, en vez de los 6000 de antes. Las unidades de caballería, y de forma correspondiente las de infantería, pasaron de tener 500 soldados a constar sólo de 120. En virtud de esta nueva ordenación, aumentó el número de legiones. Diocleciano pudo dividirlas en dos grupos: un cuerpo móvil (*comitatenses*) y tropas de frontera (*limitanei)* estacionadas de forma fija en sus acantonamientos. Para protegerse contra los jinetes persas y sus lanzas, sustituyeron la espada corta (*gladius*) con la larga (*spatha*). Así, Roma consiguió casi igualar la ventaja de equipamiento que poseían sus enemigos orientales y pudo hacer frente a la caballería sasánida. Estas tablas militares se mantuvieron durante toda la Antigüedad tardía, incluso después de que los emperadores romanos orientales —ahora desde Constantinopla, la "Nueva Roma"— asumieran la organización de la defensa de la frontera oriental. Tal equilibrio se reforzaba ocasionalmente con el pago de fuertes tributos, como los que tuvo que entregar el emperador Justiniano (quien gobernó entre 527 y 565 d. C.). Sólo en 642 d. C. acabaría la contienda entre Roma y el Imperio sasánida. Con la victoria del Islam en la región nacía una nueva época.

El autor

Theodor Kissel pertenece al Instituto de Historia Antigua de la Universidad Johannes Gutenberg de Maguncia.

Bibliografía complementaria

GRUNDZÜGE DER GESCHICHTE DES SASANIDISCHEN REICHES. Klaus Schippermann. Wissenschaftliche Buchgesellschaft; Darmstadt, 1990.

RÖMISCHE MILITÄRPOSTEN DER SEVERENZEIT AM UNTEREN CHABUR. Andreas Luther en *Göttingen Forum für Altertumswissenschaft*, n.º 5; 2002.

ROMS *LIMES* IM ORIENT. Andreas Oettel en *Schliemmans Erben. Von den Römern im Orient zur Goldstraße der Inka*. Hoffman und Campe; Hamburgo, 2003.

QREIYE/AYYASH. EIN RÖMISCHES KASTELL AM EUPHRAT. Markus Gschwind en *Zwischen Kulturen und Kontinenten*. 175 Jahre Forschung am Deutschen Archäologischen Institut. Koethen GmbH; Berlín, 2004.

La Tierra, una peonza con el corazón líquido

¿Gira la Tierra como una bola de vidrio?
No: por ser líquido su núcleo, su rotación presenta irregularidades.
Analizándolas, los geofísicos calan en el interior del planeta

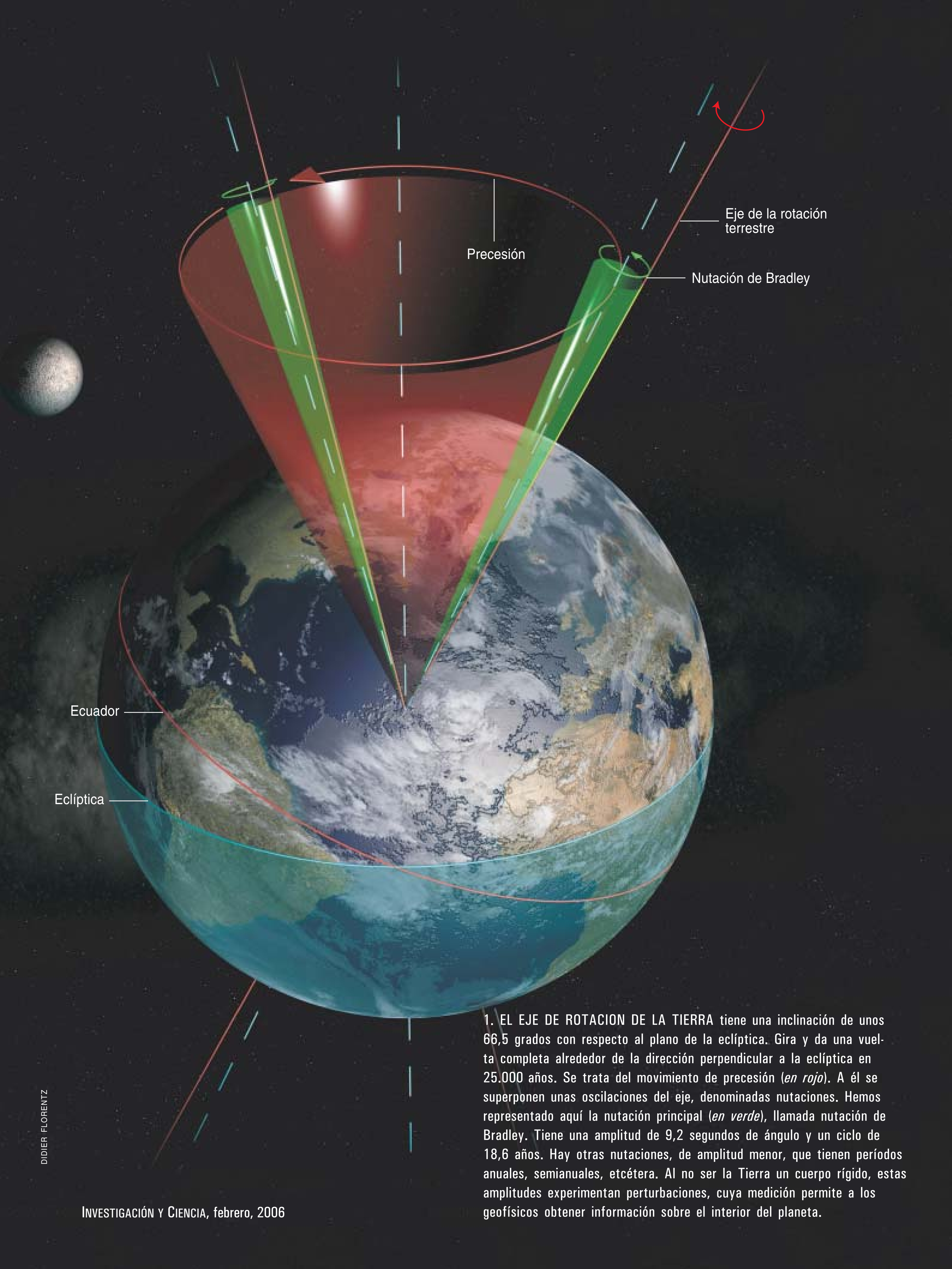
Marianne Greff-Lefftz

Tome dos huevos, uno crudo y otro duro. Hágalos girar sobre sí mismos. Inmediatamente podrá apreciar que muestran un comportamiento distinto, porque el interior del primero es fluido, mientras que el segundo es sólido.

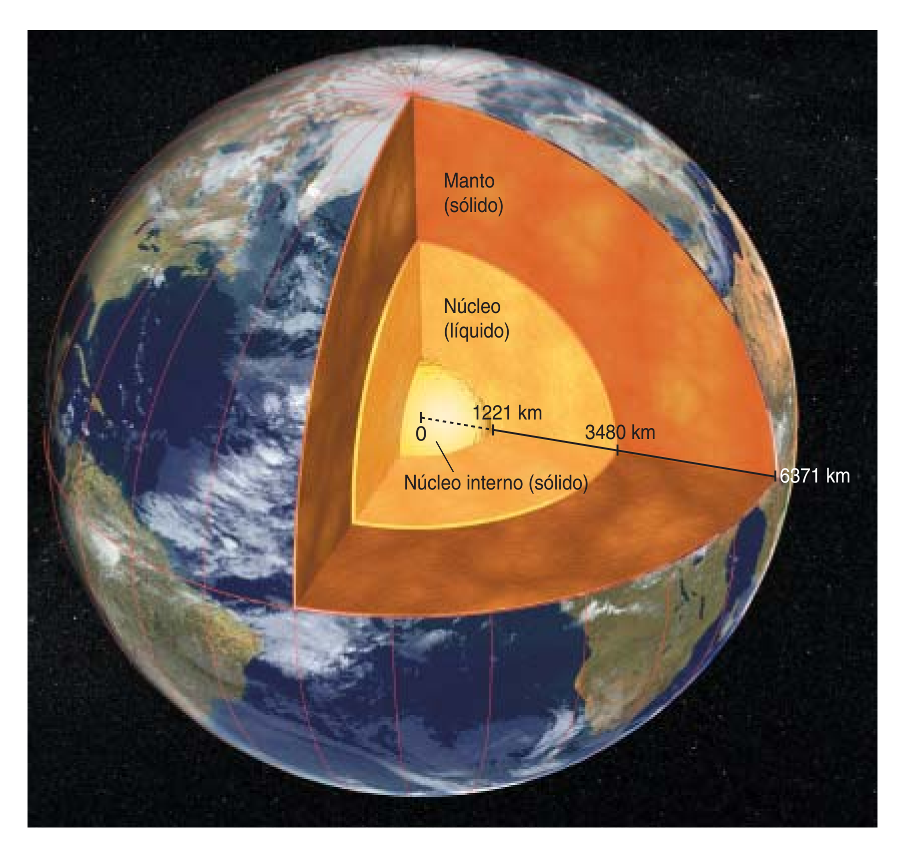
Algo similar le ocurre a la Tierra: el corazón de nuestro planeta es, en parte, líquido. La Tierra no gira del todo como lo haría una bola perfectamente sólida que tuviera su mismo tamaño, su misma forma y su misma distribución de masas. Y además su rotación está influida por la presencia de los océanos y de la atmósfera.

La rotación de la Tierra sobre sí misma constituye un movimiento de gran complejidad. En nuestros días sus variaciones se miden con mucha precisión. Desde el decenio de 1970 han ido apareciendo nuevos métodos para la observación y la medición de la posición y de la rotación terrestres, merced, sobre todo, a la técnica espacial y a los relojes atómicos. Uno de tales métodos es el sistema de localización global mediante satélites, o GPS. En otro método, llamado "interferometría de base grande" (VLBI), se mide en dos antenas distantes la diferencia entre los tiempos de recepción de las señales procedentes de una misma radiofuente extragaláctica. Un tercer instrumento, la telemetría láser-Luna, se funda en la medición del tiempo invertido por un impulso de láser en ir a la Luna y volver (gracias a reflectores que fueron instalados allí por las misiones *Apollo*). La telemetría láser-satélite, en la que el papel de la Luna lo desempeña un satélite artificial, constituye un método similar. Varias redes de estaciones u observatorios repartidos por el mundo recopilan por medio de estas diversas técnicas información sobre la rotación terrestre. Centraliza los datos obtenidos y los pone a disposición de los científicos el Servicio internacional para la rotación terrestre y los sistemas de referencia (IERS), organismo creado en 1987 cuya oficina central se encuentra actualmente en Alemania. La precisión alcanzada es muy notable. Por ejemplo, es posible rastrear el desplazamiento del polo de rotación sobre la superficie terrestre con un margen de pocos milímetros.

Para explicar las variaciones de la rotación terrestre que revelan tales medidas se construyen y analizan modelos teóricos cada vez más detallados que toman en consideración numerosos parámetros físicos, entre ellos la fluidez del núcleo terrestre. No se busca tan sólo saber reproducir la evolución de la rotación terrestre en el transcurso del tiempo, sino también detallar la estructura interna de la Tierra y los fenómenos que se desarrollan en ella. Al cotejar los resultados de los modelos teóricos con los datos procedentes de las observaciones resulta posible compilar información concerniente a parámetros geofísicos deficientemente conocidos, como las constantes asociadas al magnetismo interno, el grado de achatamiento del núcleo terrestre, la viscosidad del núcleo interno, etc. Tales informaciones sirven, a su vez, para refinar los modelos teóricos y hacer que progresen nuestros conocimientos.



1. EL EJE DE ROTACION DE LA TIERRA tiene una inclinación de unos 66,5 grados con respecto al plano de la eclíptica. Gira y da una vuelta completa alrededor de la dirección perpendicular a la eclíptica en 25.000 años. Se trata del movimiento de precesión (*en rojo*). A él se superponen unas oscilaciones del eje, denominadas nutaciones. Hemos representado aquí la nutación principal (*en verde*), llamada nutación de Bradley. Tiene una amplitud de 9,2 segundos de ángulo y un ciclo de 18,6 años. Hay otras nutaciones, de amplitud menor, que tienen períodos anuales, semianuales, etcétera. Al no ser la Tierra un cuerpo rígido, estas amplitudes experimentan perturbaciones, cuya medición permite a los geofísicos obtener información sobre el interior del planeta.



2. LA ESTRUCTURA INTERNA DE LA TIERRA se descompone en tres partes principales: el núcleo interno, el núcleo y el manto. Sobre el último descansa la corteza terrestre. El manto es sólido (pero no rígido) y está constituido por silicatos de magnesio y de hierro. El principal constituyente del núcleo es hierro metálico; se encuentra en estado líquido a causa de las elevadas temperaturas que reinan a tales profundidades. La presión aumenta con la profundidad; el hierro líquido acaba por solidificarse. La parte sólida es el núcleo interno.

¿De qué modo gira el globo alrededor de su centro de masa? Vamos a pasar revista a las características principales de la rotación terrestre: la forma en que su eje cambia de dirección con respecto a las estrellas, la forma en que cambia de dirección con respecto a la propia Tierra y cómo varía la velocidad de rotación. Y veremos lo que los geofísicos, en su esfuerzo por comprender estas variaciones, han aprendido del interior de la Tierra.

La Tierra gira sobre sí misma de oeste a este en unas 24 horas en torno a un eje norte-sur, de lo que resulta la alternancia del día y la noche. Por otra parte, el eje de rotación no es perpendicular al plano de la eclíptica —el plano de la órbita descrita por nuestro planeta alrededor del Sol—, sino que forma con ella un ángulo de unos 66,5 grados. La sucesión de las estaciones se debe a esa inclinación del eje de rotación.

Según las leyes de la mecánica, si la Tierra fuese perfectamente rígida y esférica giraría sobre sí misma, con velocidad angular constante, en torno a un eje de rotación cuya dirección permanecería fija con respecto a las estrellas. Ahora bien, nuestro planeta ni es perfectamente rígido, ni es perfectamente esférico. La rotación terrestre experimenta variaciones lo mismo en velocidad que en el eje de rotación.

El movimiento del eje de rotación terrestre se descompone en diversos movimientos de diferentes amplitudes y periodicidades. El más importante en cuanto a amplitud es el de precesión: observado desde una referencia celeste, fija con respecto al conjunto de las estrellas, el eje de rotación va girando lentamente y describe un cono de revolución cuyo eje sería perpendicular al plano de la eclíptica. El tiempo necesario para una vuelta completa es de casi 26.000 años. Este movimiento se traduce en una precesión de los equinoccios: la recta donde se cortan el plano de la eclíptica y el plano ecuatorial terrestre gira aproximadamente 50 segundos de arco al año. Este fenómeno fue observado ya por los astrónomos de la Antigüedad; se percataron de que el equinoccio de primavera (situación en la que la recta de intersección eclíptica-plano ecuatorial apunta hacia el Sol) se desplaza con respecto a las constelaciones del zodíaco. Por ese motivo, las estaciones se van desfasando con el transcurso de los siglos; al cabo de unos 13.000 años (precesión de 180 grados) las estaciones quedan invertidas.

La precesión del eje de rotación fue explicada por Newton, en el siglo XVII. Es debida a la forma ligeramente elipsoidal de la Tierra, levemente achatada por los polos y más prominente en el ecuador (la diferencia entre el radio del ecuador y el radio hasta el polo es de 22 kilómetros, para un radio medio cercano a 6370 km); este aplastamiento es resultado de las fuerzas centrífugas asociadas a la rotación. La atracción gravitatoria de la Luna y del Sol ejerce sobre la Tierra, a través del abultamiento ecuatorial, un par de fuerzas que tiende a hacer bascular su eje de rotación y a alinearlo perpendicularmente a la eclíptica. La composición de esta fuerza de basculamiento con las fuerzas centrífugas provoca una precesión. Se observa un fenómeno análogo al lanzar una peonza: su eje de rotación va girando alrededor de la vertical. En este caso, el par de fuerzas que entra en juego lo crean el peso de la peonza y la reacción ejercida por el suelo en el punto de contacto.

El movimiento del eje de rotación terrestre no es del todo uniforme, sino que, en unos pocos años, experimenta ligeras aceleraciones o deceleraciones. Asimismo, la inclinación del eje con respecto a la eclíptica oscila en torno a un valor medio. Estos movimientos, que se superponen al de precesión, se denominan nutaciones. Proceden de la variación, en el transcurso del tiempo, de las fuerzas gravitatorias ejercidas por la Luna, el Sol y los demás planetas. La nutación más importante, descubierta por el astrónomo inglés James Bradley en

DIDIER FLORENTZ

los años 1720-1740, está asociada a la evolución de la órbita lunar en torno a la Tierra (órbita que es, a su vez, perturbada por el Sol). Dicha nutación corresponde a una oscilación del eje de los polos con una amplitud de unos 9,2 segundos de arco y un período de 18,6 años. La revolución de la Tierra alrededor del Sol supone, además, nutaciones de amplitudes menores (de menos de 1 segundo de arco), de períodos anuales, semianuales, etcétera.

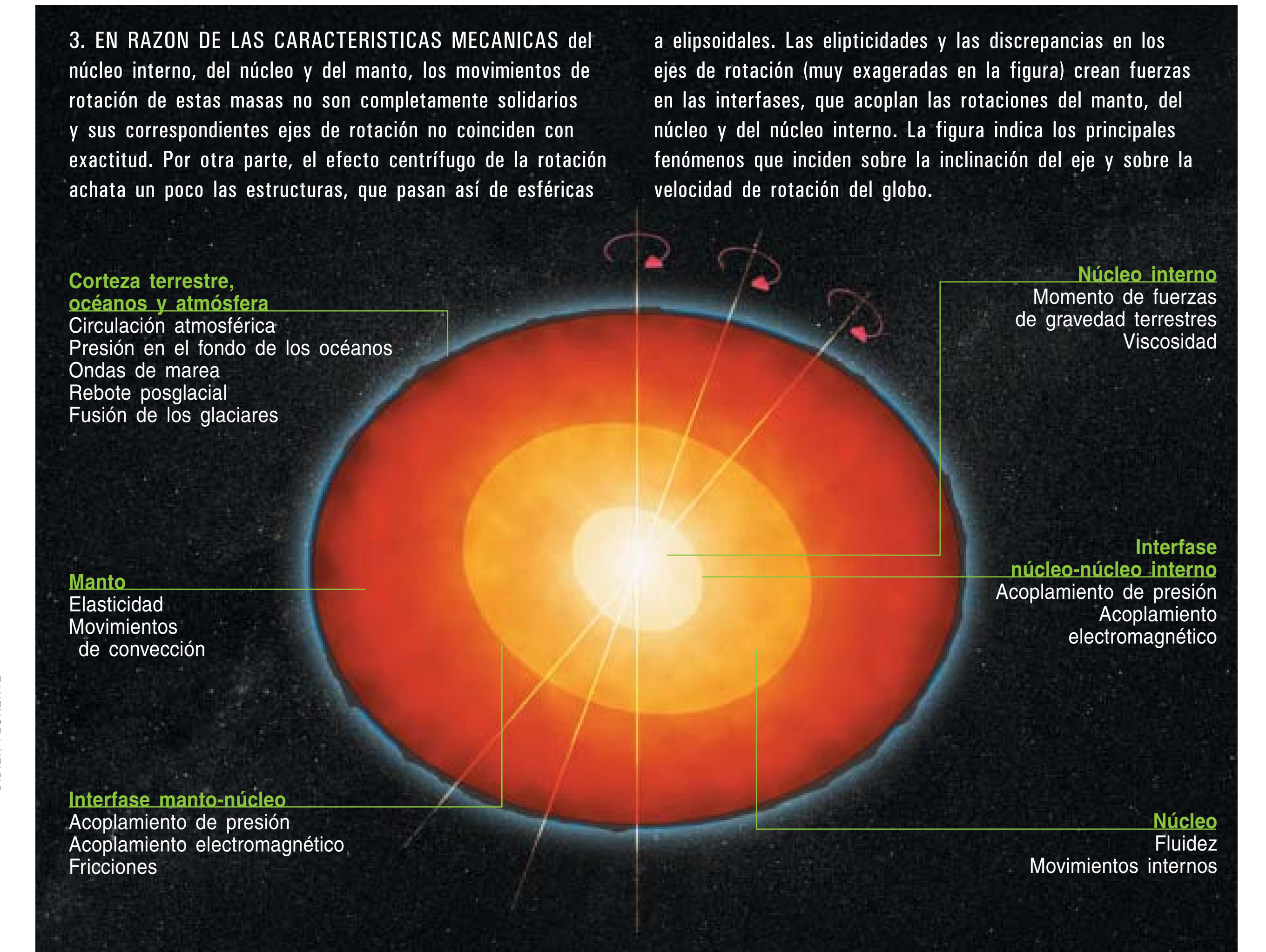
Las nutaciones y el interior de la Tierra

La precesión y las nutaciones son los movimientos clásicos de un cuerpo rígido descritos con respecto a una referencia celeste, solidaria con las estrellas. Por otra parte, al no ser la Tierra un cuerpo perfectamente rígido, la amplitud de las nutaciones experimenta perturbaciones. Los geofísicos, buscando comprender el origen exacto de estas variaciones, construyen modelos teóricos fundados en una ley conocida de la mecánica, a saber: la velocidad de variación del momento cinético (o momento angular) de un sistema es igual al momento de las fuerzas aplicadas.

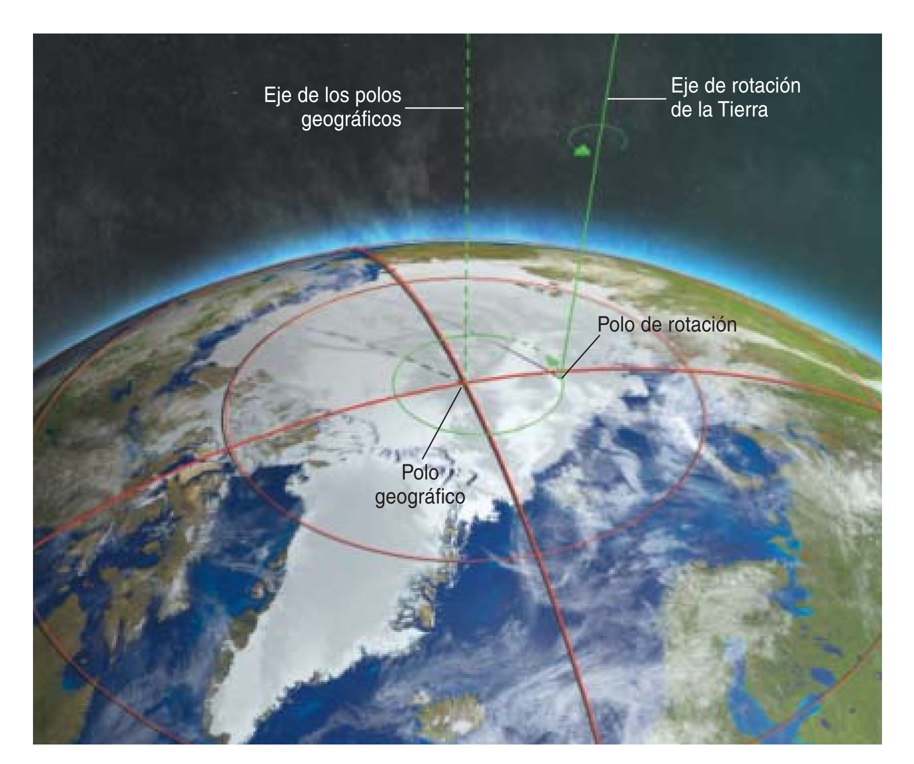
El momento cinético de un cuerpo material depende directamente de la velocidad de rotación y de su eje. Depende también de la distribución de la masa en el cuerpo y de sus eventuales movimientos internos. En cuanto al momento de las fuerzas, depende, en el caso de la Tierra —cuerpo no rígido—, de múltiples fenómenos, a menudo mal conocidos. Es necesario tener en cuenta que el interior de la Tierra está formado por un manto sólido, por un núcleo líquido y, en el centro, por una "pepita" sólida, el núcleo interno. Los movimientos de rotación de estas tres entidades, aunque acoplados, no coinciden. Tres momentos de fuerzas internas revelan ser importantes para explicar las variaciones de la rotación terrestre: un momento de fuerzas de presión, un momento de fuerzas gravitatorias y un momento de fuerzas electromagnéticas. El primero de ellos se conoce desde hace casi un siglo. Los otros dos no se identificaron hasta los años noventa, por John Wahr, de la Universidad de Colorado, Sonny Mathews, de la Universidad de Madrás, Bruce Buffet, de la Universidad de Columbia Británica, y Thomas Herring, del Instituto de Tecnología de Massachusetts.

El momento debido a las fuerzas de presión, denominado par inercial, resulta de la rotación en bloque del núcleo fluido. El matemático francés Henri Poincaré, considerando una Tierra elipsoidal en rotación y sometida a una precesión, demostró en 1910 que si el núcleo terrestre estuviera constituido por un fluido perfecto, es decir, un líquido carente de viscosidad, entonces giraría como un bloque alrededor de un eje cuya dirección se hallara levemente inclinada con respecto al eje de rotación terrestre (del manto). Resulta así que el manto y el núcleo no giran exactamente alrededor del mismo eje (según los cálculos actuales, se estima que, en la superficie del núcleo, la separación entre los dos ejes es de unos 30 metros). Ahora bien, el núcleo, al igual que el manto, está levemente aplastado y, al girar, "tropieza" con el manto, creando así presiones en la interfase manto-núcleo cuyo momento resultante no es nulo. Señalemos que si el núcleo fuese perfectamente esférico giraría sin chocar contra el manto y el correspondiente par inercial sería

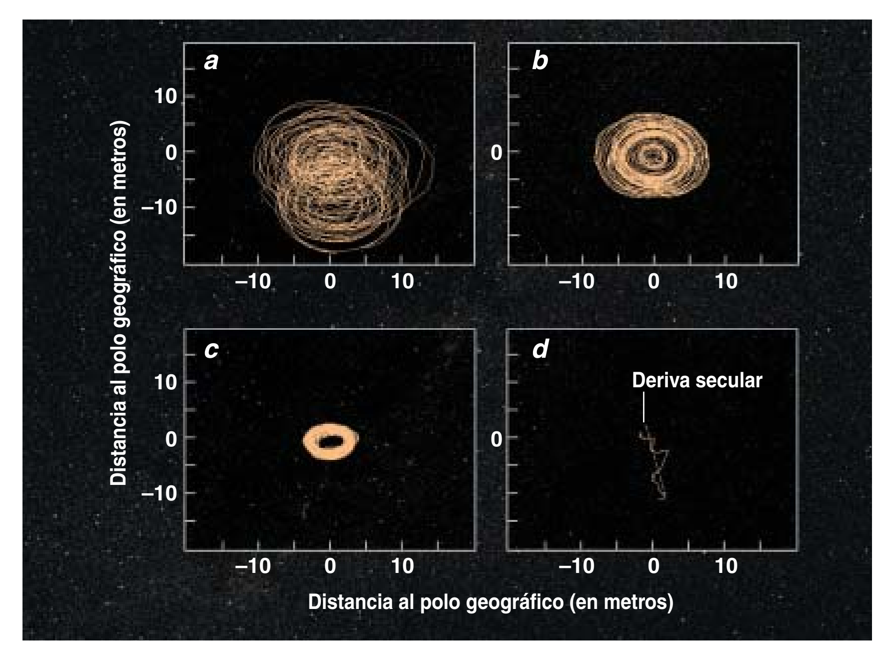
3. EN RAZON DE LAS CARACTERISTICAS MECANICAS del núcleo interno, del núcleo y del manto, los movimientos de rotación de estas masas no son completamente solidarios y sus correspondientes ejes de rotación no coinciden con exactitud. Por otra parte, el efecto centrífugo de la rotación achata un poco las estructuras, que pasan así de esféricas a elipsoidales. Las elipticidades y las discrepancias en los ejes de rotación (muy exageradas en la figura) crean fuerzas en las interfases, que acoplan las rotaciones del manto, del núcleo y del núcleo interno. La figura indica los principales fenómenos que inciden sobre la inclinación del eje y sobre la velocidad de rotación del globo.



DIDIER FLORENTZ



4. EL POLO GEOGRAFICO no coincide exactamente con el polo de rotación, lugar donde el eje de rotación instantáneo atraviesa la superficie de la Tierra. En un sistema de referencia solidario con la Tierra, el polo de rotación describe una trayectoria aproximadamente circular durante un período de unos 14 meses, con un radio de algunos metros (en la figura, la distancia entre el polo norte geográfico y el polo norte de rotación está muy exagerada).



5. SE HA REPRESENTADO en *a* el desplazamiento con respecto al suelo del polo de rotación desde 1900. El cero del eje de abscisas corresponde al meridiano de Greenwich, mientras que el cero del eje de ordenadas corresponde al meridiano 90º este; el punto de coordenadas (0,0) corresponde al polo geográfico. El desplazamiento se descompone en tres movimientos: la oscilación de Chandler (*b*), con un período de 14 meses, una oscilación forzada anual (*c*) y una deriva secular combinada con variaciones decenales (*d*). Los trazados han sido realizados a partir de datos facilitados por el IERS.

nulo. Idénticas consideraciones son aplicables para la frontera existente entre el núcleo, fluido, y el núcleo interno, que es sólido.

El segundo momento de fuerzas, que es de naturaleza gravitatoria, se debe a la forma elipsoidal del manto, del núcleo y la "pepita" del núcleo terrestres, y al hecho de que los ejes de rotación de estos tres elementos son ligeramente distintos. Se demuestra que las masas del manto y del núcleo ejercen sobre el núcleo interno (la "pepita") fuerzas gravitatorias cuyo momento resultante no es nulo. La autora, en colaboración con Véronique Dehant, del Real Observatorio de Bélgica, y con Hilaire Legros, del Instituto de Física del Globo de Estrasburgo, ha calculado recientemente las consecuencias de este momento gravitatorio sobre las nutaciones. Dicho momento desempeña un papel clave. Depende estrictamente del comportamiento reológico del núcleo interno, es decir, de su capacidad de ser deformado bajo la acción de una fuerza. Si el núcleo interno es rígido o elástico, el momento gravitatorio tenderá a bloquearlo con respecto al manto. Por el contrario, si es viscoso, tenderá a fluir, a "ceder", y este momento acaba por anularse.

El tercer momento de las fuerzas internas que intervienen en las nutaciones es de origen electromagnético. El manto inferior, el núcleo y el núcleo interno están constituidos por materiales conductores de la electricidad; esos tres elementos giran de formas ligeramente diferentes. En las interfases manto-núcleo y núcleo-núcleo interno, las discrepancias en los movimientos de rotación de la materia en el seno de los campos magnéticos que allí existen provocan la inducción de corrientes eléctricas. Ahora bien, en presencia de un campo magnético, un conductor recorrido por una corriente eléctrica experimenta una fuerza. Resultan de ello pares de fuerzas electromagnéticas que tienden a impedir la aparición de diferencias excesivamente grandes en las velocidades de rotación del manto, del núcleo y del núcleo interno.

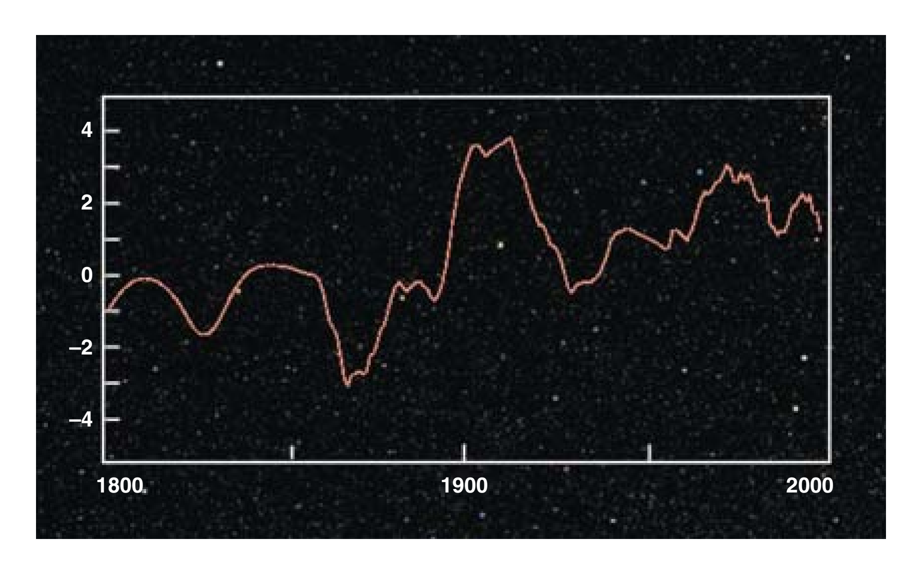
Tomados en consideración estos tres momentos de fuerzas internas, las ecuaciones que gobiernan los movimientos de los ejes de rotación de la Tierra (es decir, del manto, del núcleo y del núcleo interno) hacen

DIDIER FLORENTZ (*arriba*)

aparecer cuatro frecuencias propias; dicho de otro modo, cuatro frecuencias de movimientos propios del sistema manto-núcleo-núcleo interno) y ello, independientemente de las acciones externas (que serían las fuerzas gravitatorias del Sol y de la Luna). Una primera frecuencia propia es consecuencia de que el eje de rotación de la Tierra no coincide con el eje de simetría del geoide (sobre lo cual volveremos).

El ciclo es de 435 días en la referencia terrestre, es decir, como lo percibiría un observador que girase solidariamente con la Tierra. A ella se añaden dos frecuencias propias cuasidiurnas (su período es muy próximo a 24 horas, también en el sistema referencial terrestre). Una de ellas está asociada a la existencia de un núcleo líquido: es la denominada frecuencia de nutación libre del núcleo, (FCN, de *Free Core Nutation*). La otra está vinculada con el basculamiento del núcleo interno con respecto al núcleo. Es la frecuencia de nutación libre del núcleo interno (FICN, de *Free Inner Core Nutation*). La cuarta frecuencia propia atañe al basculamiento del núcleo interno con respecto al manto. El período correspondiente es de 6,3 años.

Las frecuencias propias, intrínsecas al sistema considerado, no siempre se someten a una observación directa. Una estructura mecánica —por ejemplo, un puente— posee una frecuencia de vibración propia, o frecuencia "natural" de resonancia, que se manifiesta tras una excitación externa adecuada. Cuando se hace que el puente vibre a una frecuencia próxima a la frecuencia propia, se produce un fenómeno de resonancia y la vibración se amplifica. De igual modo, los fenómenos de resonancia ponen de manifiesto las frecuencias propias de la Tierra. Tal es el caso de las frecuencias cuasidiurnas de nutación libre del núcleo y del núcleo interno. Observadas en sistema de referencia terrestre, ciertas fuerzas gravitatorias (con mayor precisión, las ondas de marea) ejercidas por la Luna y el Sol varían con un período casi igual a un día. Se produce entonces una resonancia con las dos frecuencias propias de nutación libre. El resultado es una amplificación de las nutaciones con respecto a las que hubiera tenido



6. ESTA GRAFICA INDICA LA DESVIACION, medida en milésimas de segundo, de la media anual de la duración del día con respecto a las 24 horas exactas en el período de 1800 a 2000. Tomando un siglo como unidad de escala, el aumento medio de la duración del día es del orden de dos milisegundos.

SEGUN DATOS DEL *IERS*

una Tierra rígida, falta de núcleo y de núcleo interno.

La frecuencia de nutación libre del núcleo se puso así de manifiesto en las mediciones de la nutación del eje de rotación terrestre efectuadas en los años ochenta. Su valor discrepa en más del 5 por ciento del valor teórico calculado algunos años antes. Este resultado indica un achatamiento del núcleo de más del 5 por ciento con respecto al supuesto en los modelos, los cuales tan sólo contemplaban el aplastamiento debido al equilibrio entre la fuerza centrífuga de la rotación y la autogravitación de la Tierra. Este achatamiento adicional está vinculado, probablemente, a heterogeneidades de la densidad en el seno del manto terrestre, debidas, a su vez, a movimientos de convección. Por otra parte, las mediciones demuestran que la componente de nutación libre del núcleo se amortigua en el transcurso del tiempo. Los geofísicos deducen de aquí la existencia de importantes fricciones entre el núcleo y el manto.

En el año 2000, tras un análisis de las amplitudes de nutaciones anuales, obtenidas mediante interferometría de base muy grande, Mathews y su grupo propusieron un valor para la frecuencia de nutación libre del núcleo interno. Dicho valor difiere considerablemente del valor teórico, pero éste no toma en consideración los fenómenos de fricción magnética entre el núcleo interno y el núcleo. Un análisis de las nutaciones, extendido a períodos más prolongados, debería permitir a los geofísicos la determinación precisa del período y del amortiguamiento de la nutación a la frecuencia de nutación libre del núcleo interno, lo cual informará de la amplitud del campo magnético en la interfase entre el núcleo y el núcleo interno y de la viscosidad de este último.

El polo de rotación se desplaza

La precesión y la nutación del eje de rotación de la Tierra constituyen movimientos detectables desde una referencia celeste. Existe otro aspecto de la rotación terrestre que se manifiesta si referimos el eje de rotación a la Tierra misma, localizando el punto del suelo en el cual el eje de rotación atraviesa a la superficie de la Tierra. Contrariamente a la creencia general, el eje de rotación de la Tierra no coincide con el polo geográfico. Los polos norte y sur corresponden al eje de simetría del elipsoide terrestre. El eje de rotación no se confunde con este eje de simetría, ni permanece fijo con respecto al suelo terrestre.

Los astrónomos descubrieron ya a finales del siglo XIX que el polo de rotación describe en la superficie de la Tierra una trayectoria irregular, aproximadamente circular, inscrita en un cuadrado de una veintena de metros de lado, que contiene el polo

geográfico. El polo de rotación describe una vuelta completa en unos 14 meses. Fenómenos de este tipo habían sido predichos ya para cuerpos sólidos en rotación en el siglo XVIII. Leonhard Euler, considerando un elipsoide en rotación, había establecido que si, por una razón cualquiera, el eje de rotación no estuviera alineado con el eje de simetría del elipsoide, el primero de estos ejes efectuaría un movimiento de precesión alrededor del segundo, con una frecuencia calculable. Indiquemos que se trata de una frecuencia propia, que podría ser revelada mediante una excitación externa adecuada. Al aplicar el cálculo euleriano a una Tierra elipsoidal y perfectamente rígida se encuentra que el eje de rotación debería dar una vuelta en torno al eje de simetría cada 298 días. Dicho de otro modo, veríamos al polo de rotación describir un círculo alrededor del polo cada diez meses.

La trayectoria que realmente describe el polo de rotación es más complicada que la predicha por el cálculo de Euler. Cabe descomponerla en tres movimientos principales: oscilación de Chandler, oscilación anual forzada y de una desviación de la posición media del polo.

La oscilación de Chandler se llama así en recuerdo del astrónomo norteamericano Seth Chandler, que descubrió en 1891 que el polo de rotación se desplaza con respecto al suelo terrestre. Esta oscilación, cuya amplitud tan pronto se ve amortiguada como resulta amplificada, tiene un período de 435 días (*véase la figura 5b*). Corresponde al fenómeno revelado por Euler, a pesar de que su período no sea de 298 días: la elasticidad del manto, así como la presencia de un núcleo fluido y de océanos, modifican la frecuencia propia. De acuerdo con trabajos publicados en 2000 por el Laboratorio de Propulsión a Chorro de California, la excitación es debida sobre todo a variaciones de presión en el fondo de los océanos (provocadas por cambios de temperatura, de salinidad o por corrientes), a fluctuaciones de la presión atmosférica y a los vientos.

La segunda componente del movimiento polar consiste en una oscilación anual forzada (*véase la figura 5c*). Es debida a la circulación atmosférica, que hace bascular a la Tierra con un ritmo marcado por las estaciones.

La tercera componente consiste en una deriva de la posición media del polo (*véase la figura 5d*), que se descompone, a su vez, en dos movimientos. Una deriva de tendencia secular, del orden de algunos centímetros por año en dirección a Canadá es consecuencia probable de movimientos de convección en el manto terrestre, de la fusión de glaciares y del "rebote posglacial" (la fusión de los casquetes polares a partir de la última glaciación ha implicado una descompresión del manto viscoso, de suerte que la superficie actual se eleva aproximadamente entre 20 y 50 centímetros por siglo). El segundo movimiento de deriva, más irregular, consiste en variaciones decenales de origen controvertido.

La duración del día fluctúa

La dinámica de la rotación de la Tierra se traduce así en los movimientos de su eje, ya se trate de su descripción en una referencia celeste (precesión y nutaciones) o de una referencia terrestre (desplazamientos del polo de rotación con respecto al suelo). También tiene traducción en variaciones de la velocidad de rotación en torno a dicho eje, fenómeno que modifica la duración del día. Gracias al estudio de los eclipses de la Antigüedad y a determinadas observaciones geológicas, se sabe que la duración del día va aumentando poco a poco. Y desde hace varios decenios han podido ser medidas numerosas irregularidades en esta duración, merced a la introducción de relojes de alta precisión, como los atómicos.

La duración del día aumenta del orden de dos milisegundos por siglo. La causa de esta minúscula lentificación en la velocidad de giro es la fricción en el fondo de los océanos debida a las mareas, que disipa parte de la energía cinética de rotación de la Tierra. A esta infinitesimal ralentización se suman varias perturbaciones periódicas, cuyas amplitudes son del orden de la milésima de segundo. Los movimientos de la atmósfera terrestre provocan oscilaciones en la duración del día con períodos inferiores a dos años. El rebote glacial y la fusión de los glaciares en la actualidad son responsables de variaciones seculares. Por otra parte, numerosos estudios han demostrado que la influencia del núcleo sobre la duración del día se traduce en variaciones decenales. Para comprenderlas bien, los geofísicos deberían determinar el acoplamiento entre el momento cinético del manto y el correspondiente al núcleo, un acoplamiento en el que intervienen numerosos fenómenos (efectos deviscosidad y de presión, efectos electromagnéticos y gravitatorios). Al igual que los movimientos del eje, las variaciones de la velocidad de rotación dependen en parte de lo que acontece en las profundidades de la Tierra.

Los geofísicos han logrado así, por medio del estudio del eje y de la velocidad de rotación de la Tierra, conocer muchos detalles sobre los achatamientos del núcleo y del núcleo interno, sobre la amplitud del campo magnético en la interfase núcleo-manto, sobre las fricciones de origen electromagnético entre el núcleo y el manto o sobre la viscosidad del núcleo interno. A pesar de todo ello, nuestro conocimiento del interior de la Tierra sigue siendo insuficiente. Para perfeccionarlo, así como para comprender mejor la rotación de la Tierra, los geofísicos habrán de seguir confrontando los resultados de sus modelos teóricos y los registros de la rotación terrestre, sin olvidar los experimentos de laboratorio y los métodos sismológicos.

La autora

Marianne Greff-Lefftz es geofísica del Instituto de Física del Globo de París (IPGP).

Bibliografía complementaria

EFFECTS OF INNER CORE VISCOSITY ON GRAVITY CHANGES AND SPATIAL NUTATIONS INDUCED BY LUNI-SOLAR TIDES. M. Greff-Lefftz, V. Dehante y H. Legros en *Physics of the Earth and Planetary Interiors*, vol. 129, págs. 31-41; 2002.

MODELLING OF NUTATION AND PRECESSION: NEW NUTATION SERIES FOR NONRIGID EARTH AND INSIGHTS INTO THE EARTH'S INTERIOR. P. Mathews, T. Herring y B. Buffett en *Journal of Geophysical Research (Solid Earth)*, vol. 107; 2002.

Marc Boada

Levitación magnética

Cuenta Plinio el Viejo en su *Historia Natural* que Dinócrates, arquitecto de Alejandría, trazó para la reina Arsinoes un plano para la construcción de la bóveda de un templo en piedra imán, de suerte que una escultura labrada en hierro pareciese pender en el aire. San Agustín lo dio por cierto. Mediante el cálculo de la fuerza de los imanes, explica, los sacerdotes del templo consiguieron que la estatua de hierro se mantuviera en perfecto equilibrio, haciendo creer a los fieles que se trataba de un milagro perpetuo. Si bien la levitación magnética exige algo más que unos bloques de magnetita, el ejemplo muestra que los imanes vienen cautivando a las mentes inquietas desde la antigüedad.

La magnetita (Fe_30_4) constituye un mineral de color negro, cristalino o masivo, que, de forma natural, atrae al hierro. Hasta el advenimiento de la electricidad, fue el único imán permanente con el que se podía experimentar. Sus propiedades permanecieron envueltas en el misterio durante siglos.

Hoy se acepta que un imán permanente corresponde a un cuerpo en donde los momentos magnéticos de los electrones están orientados en paralelo, es decir, los espines electrónicos están dirigidos en el mismo sentido. Ello determina la aparición, en el cuerpo imanado, de dos polos, norte y sur, a partir de los cuales se extiende un campo magnético.

Quizá Plinio desconocía que ni siquiera la magnetita de la mejor calidad podía ejercer la fuerza necesaria para hacer levitar una escultura de hierro. En cualquier caso, este material, manipulado de forma adecuada, aumenta notoriamente su fuerza magnética.

Siguiendo los métodos tradicionales, tomemos un trozo de magnetita que, de forma espontánea, atraiga clips de oficina. Tamicemos sobre ella finas virutas de hierro. Comprobaremos que presenta diversos polos. De ellos, escogeremos, mediante una brújula, dos de signo opuesto y situados aproximadamente uno en las antípodas del otro. Con cuidado e intentando que la pieza no se caliente, tallaremos, mediante una sierra de diamante, dos facetas paralelas orientadas perpendicularmente a un eje imaginario que une los dos polos seleccionados.

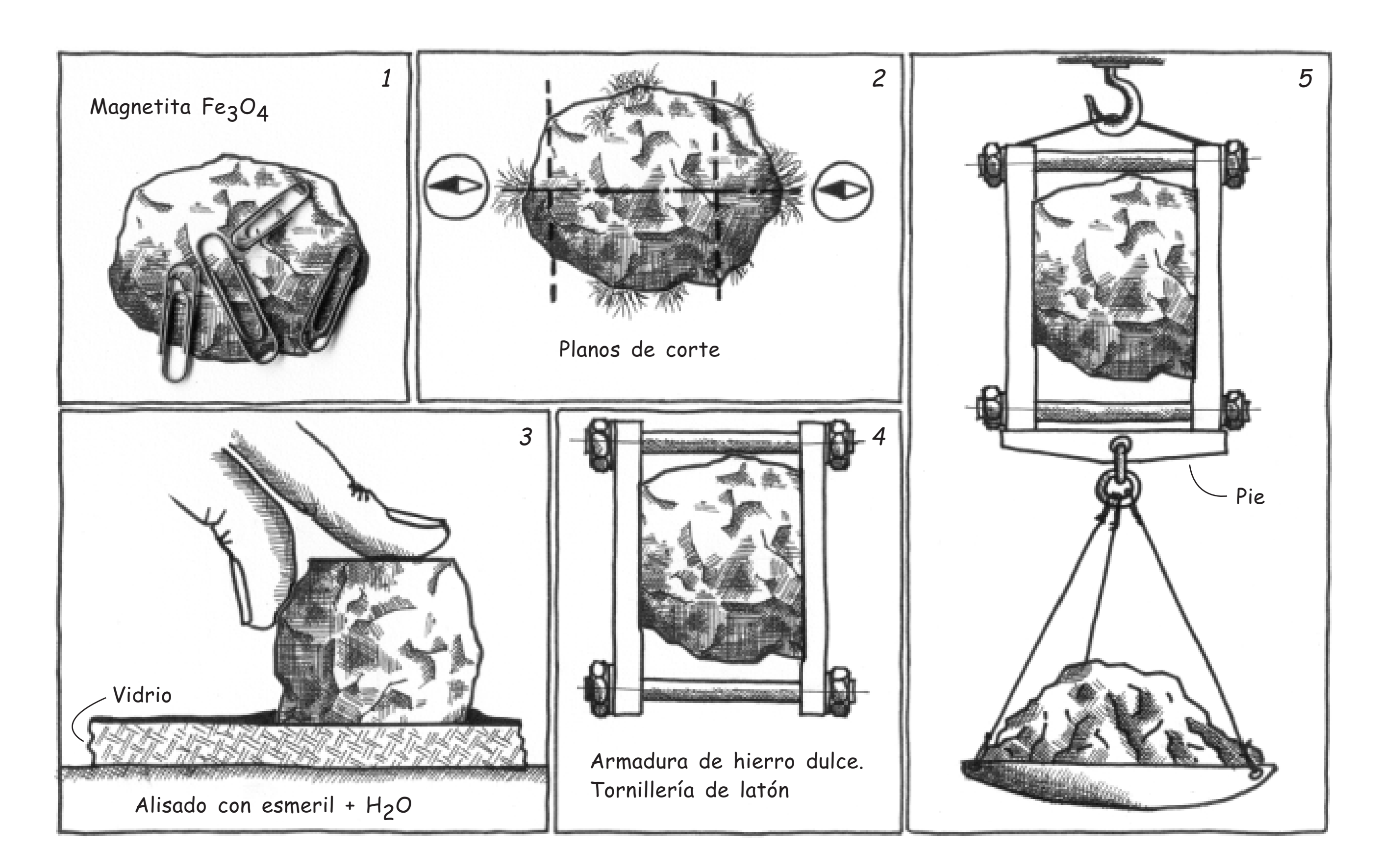
Alisaremos las facetas por abrasión sobre un cristal con la interposición de algún abrasivo en grano ligeramente húmedo. Construiremos luego un emparedado situando la magnetita entre dos placas de hierro dulce (libre de impurezas), que atornillaremos mediante espigas y tuercas de latón. Estas zapatas polares en contacto íntimo con la magnetita canalizan el campo magnético y lo concentran en sus extremos, donde instalaremos una tercera pieza, antiguamente llamada pie, que cerrará éste.

Estas manipulaciones se realizan con el fin de reforzar el campo magnético del mineral y observar así la aparición de un curioso fenómeno muy conocido. Si al pie que une las zapatas polares acoplamos un recipiente liviano y a éste añadimos peso de forma progresiva (mediante perdigones, por ejemplo), incrementaremos poco a poco el campo magnético de la magnetita hasta que ésta soportará masas equiparables a la suya. (Este fenómeno se observa en artefactos históricos de gran tamaño, como el expuesto en el Museo Nacional de Ciencias Naturales de Madrid o el del Museo de Historia de la Ciencia de Florencia.)

La imanación no aumenta de forma gradual, sino a saltos. Ello se debe a que en los imanes naturales la


MARC BOADA

Levitador magnético construido con dos imanes de neodimio y frenos magnéticos de cobre.



estructura cristalina es muy heterogénea: presenta dominios magnéticos (o recintos de Weiss) que se alinean "de golpe" con el campo magnético principal. Este aumento brusco del campo se conoce como efecto Barkhausen. En la magnetita resulta notable precisamente por el considerable tamaño de los recintos de Weiss iniciales.

Como hemos dicho, el lento incremento de peso que soporta nuestro imán conduce a un aumento en la intensidad de su campo magnético, que, a su vez, induce una mejor orientación de los recintos. Esta realimentación prosigue hasta alcanzar un límite en que el peso del pie, junto con el de la masa que soporta, resulta excesivo; por ello, cae y destruye el campo magnético. Verificamos entonces que la magnetita no soporta tanto peso, por lo que debemos reiniciar el proceso. En sus *Elementos de Física Experimental*, de 1841, Claude Pouillet lo advierte con claridad: "con tiempo y precaución se llegará a nutrirle de nuevo y a devolverle su primer vigor".

¿A qué responde este comportamiento? En primer lugar, el aumento a saltos de la imanación puede deberse no sólo a un campo inductor de la orientación, sino también a tensiones internas del material (el acero se magnetiza débilmente martilleándolo en frío) y otros factores mecánicos, sino también a fluctuaciones graduales de la temperatura (el acero se magnetiza calentándolo al rojo cereza y dejándolo enfriar a ser posible en posición vertical). En segundo lugar, un imán que soporta un peso en aumento realiza un trabajo que

IMANES PERMANENTES

Los imanes permanentes se construyen con materiales ferromagnéticos en los que existen pequeños dominios con la misma orientación de los dipolos magnéticos elementales. Se denominan recintos de Weiss. Su tamaño oscila entre algunos centímetros y pocos milímetros. En su interior los espines de los electrones libres se disponen en paralelo, apuntando todos en el mismo sentido.

En estado no imanado, los recintos adoptan orientaciones desordenadas, de forma que su efecto global es nulo. Están separados por las paredes de Bloch, delgadas zonas de transición de pocas decenas de átomos de espesor. El comportamiento de éstas guarda semejanza con el de la superficie de un líquido, donde la tensión superficial tiende a reducir su área para llevarla a un estado de mayor estabilidad termodinámica.

Cuando se magnetiza un imán, los recintos que forman un ángulo obtuso con la dirección del campo exterior experimentan giros bruscos y de gran amplitud (efecto Barkhausen); en general, se trata de un proceso irreversible. Los recintos que forman un ángulo agudo con el campo exterior, en cambio, giran sólo unos pocos grados; cuando éste cesa, pueden volver a su posición inicial.

Dominios magnéticos en un material ferromagnético. En cada uno de ellos, los momentos magnéticos de los electrones se orientan en paralelo.

requiere cierta energía. Recordemos que un imán con todos los dominios alineados es termodinámicamente más estable que otro con una distribución contraria. A expensas de la diferencia de energía entre los estados desordenados de los recintos de Weiss se realiza el trabajo de sustentación del peso creciente.

Por fin, cuando el peso cae, la magnetita pierde gran parte de su magnetismo. Ello se debe a las dislocaciones del retículo cristalino: los recintos de Weiss que no se habían alineado ayudan a que otros vuelvan a su posición inicial, una vez ha cesado el trabajo de sustentación.

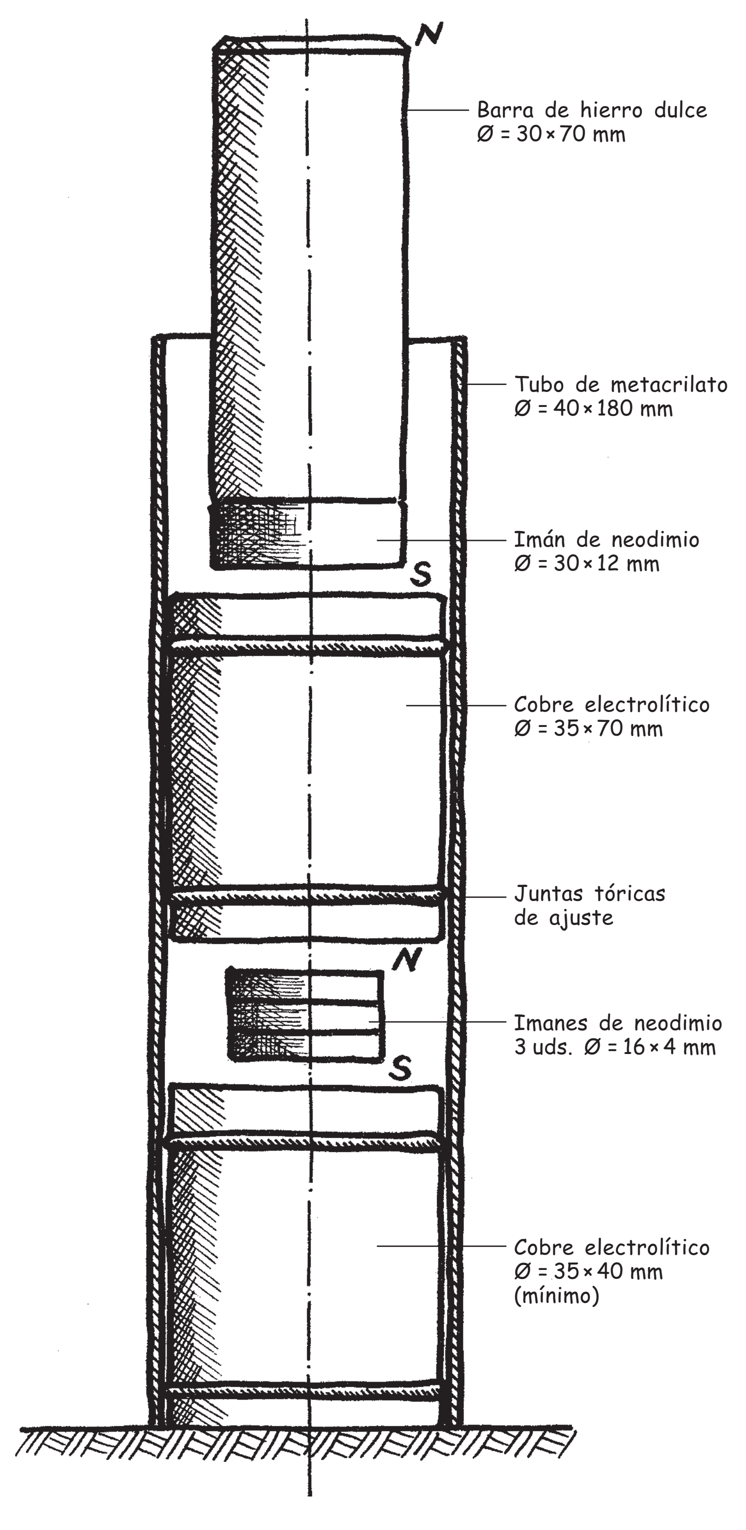
En la actualidad disponemos de imanes cerámicos con campos magnéticos de intensidad miles de veces superior al de la magnetita natural. La fabricación por sinterizado permite ajustar con precisión el tamaño de los dominios magnéticos y los tratamientos térmicos. Merced al conocimiento de los ciclos de imanación, se obtienen materiales que permiten hacer realidad, en parte, el sueño de Arsinoe.

Con mi colaborador Steve Gennrich, del Exploratorium de San Francisco, hemos construido un levitador magnético de gran simplicidad técnica donde es el propio experimentador quien aporta el control y la energía necesarios para el experimento. El prototipo forma parte del proyecto de museo del cobre *La Farga Lacambra*. Se ha desarrollado merced a la colaboración de Joaquim Pla, M.ª Àngels Crusellas y Miquel Caballeria, de la Universidad de Vich, y a la financiación de la entidad promotora de dicho museo, la Corporació Metal·lúrgica Catalana, S. L. En esencia, consiste en dos imanes de neodimio, uno superior ligeramente mayor que el inferior. Las fuerzas de atracción que operan entre polos opuestos se encargan de que el imán inferior sea atraído por el de arriba, ascienda y empiece a flotar en el aire.

Para facilitar la manipulación, utilizamos dos frenos magnéticos. Constan éstos de sendos discos de cobre, donde los imanes de neodimio, gracias a su potente campo magnético, inducen corrientes eléctricas con campos magnéticos asociados de sentido opuesto al del campo inductor. Se trata de las corrientes eléctricas de Foucault. Se perciben con facilidad cuando se mueve un imán cerámico sobre un conductor de espesor suficiente. Así, tomando el imán con sus dedos, el experimentador notará una fuerte resistencia al avance; un tacto viscoso. Ello se debe a que, según la regla de Lenz, el campo magnético propio de estas corrientes se halla orientado de tal forma, que se opone al movimiento del campo que las induce.

Las dimensiones de los discos de cobre deben determinarse empíricamente, sobre todo el que se encuentra entre los dos imanes. Se exige que sea delgado para que el imán superior atraiga al inferior. Al disco de la base se le da, como mínimo, el mismo espesor que al anterior.

Determinaremos luego el espacio entre los elementos de cobre donde el imán inferior levitará. Para ello introduciremos entre éstos pequeñas columnas de material no magnético (*no se muestran en la ilustración*) con una altura similar al diámetro del imán levitante. Por fin, añadiremos al imán superior una barra de hierro dulce, que alejará el polo superior de la zona de experimentación, facilitando además su asimiento.



Llegados a este punto, iniciaremos el experimento situando los imanes en su posición, con los polos opuestos enfrentados. Acercaremos desde arriba y lentamente el imán superior; observaremos que el inferior empieza a elevarse. La práctica permite mantenerlo durante tiempo casi estático entre los discos de cobre. Para ello debe moverse ligeramente el imán superior. En ausencia de movimiento las corrientes de Foucault desaparecen, quedando el sistema sin el freno magnético que, mediante la ralentización de los movimientos, facilita el control fino de la distancia entre los imanes.

Juan M.R. Parrondo

Incentivar la sinceridad

Alicia es presidente de una pequeña comunidad de cuatro vecinos: ella misma, Berta, Carlos y Daniel. Alicia pretende construir un ascensor en el edificio. El coste total asciende a 10.000 euros, pero no parece justo que todos paguen lo mismo. Algunos tienen una mayor necesidad del ascensor, como los que viven en los últimos pisos.

Alicia intenta en primer lugar un método bastante ingenuo. En una de las juntas pregunta públicamente a cada vecino (incluyéndose a sí misma en la encuesta) cuánto estaría dispuesto a pagar para tener el ascensor. Enseguida se da cuenta de que todos tienden a decir una cifra más baja de la que realmente estarían dispuestos a pagar, para ahorrarse algún dinero o por miedo a que los otros hagan lo mismo.

¿Cómo hacer que todos digan la verdad? Existe una ingeniosa solución del problema, aunque no completamente satisfactoria. Se trata del llamado mecanismo de Vickrey-Clarke-Groves (VCG), inspirado en las subastas de Vickrey —que abordamos aquí el mes pasado— y diseñado por los economistas norteamericanos Edward Clarke y Theodore Groves.

La idea es la siguiente. Alicia anuncia que va a preguntar a cada vecino lo que está dispuesto a pagar y que anotará las cantidades ofrecidas por cada uno, que llamaremos d_A, d_B, d_C y d_D. Si la suma de las cuatro ofertas alcanza los 10.000 euros, el ascensor se instalará. Hasta ahora el sistema es idéntico al primer intento ingenuo de Alicia. Lo que cambia en el mecanismo VCG es la contribución de cada vecino, que no pagará la cifra que ha declarado o ni siquiera una cantidad proporcional. Alicia pagará una cantidad dada por la siguiente fórmula:

$$c_A = 10.000 - (d_B + d_C + d_D).$$

Es decir, lo que paga Alicia no depende de lo que ella declara, sino sólo de lo que han ofrecido sus vecinos.

Esta extraña regla tiene una gran virtud: induce a decir la verdad. Supongamos que Alicia está dispuesta a pagar una cantidad v_A para tener el ascensor. Es decir, el valor real del ascensor para Alicia es exactamente v_A: si tuviera que pagar más estaría perdiendo dinero y si tuviera que pagar menos estaría ganando. (El mes pasado, al hablar de las subastas de Vickrey, comentamos con algo más de detalle el significado de este valor o "utilidad" de un bien, un concepto clave en economía.) Con más precisión, en caso de que el ascensor se instalara, Alicia "ganaría" una cantidad (que puede ser positiva o negativa):

$$U_A = v_A - c_A$$

¿Qué suma d_A debe ofrecer Alicia para maximizar su ganancia U_A? Ella no conoce el valor v_B, v_C, v_D, que tiene el ascensor para el resto de inquilinos, ni tampoco lo que éstos van a declarar. Pero sabe que el ascensor se instalará si el total de ofertas supera los 10.000 euros, es decir, si:

$$d_A > 10.000 - (d_B + d_C + d_D) = c_A.$$

Habrá, pues, ascensor si la oferta d_A de Alicia supera la cantidad c_A que ella misma pagará en caso de que la obra se acometa.

Si Alicia ofrece una suma d_A mayor que el valor real v_A, puede ocurrir que el resto de ofertas sea tal, que $d_A > c_A > v_A$. En este caso el ascensor se instalará y Alicia perderá dinero. Si, por el contrario, ofrece una suma d_A menor que el valor real v_A, puede ocurrir que el resto de ofertas sea tal que $d_A < c_A < v_A$, con lo que el ascensor no se construiría a pesar de que, si se hiciera, Alicia ganaría una cantidad U_A positiva, porque su contribución c_A sería inferior al valor v_A que el ascensor tiene para ella. Por lo tanto, la decisión óptima de Alicia es decir la verdad: declarar exactamente el valor real v_A. Lo mismo ocurre para el resto de vecinos.

El mecanismo VCG está inteligentemente diseñado para que mentir sea una mala estrategia. Si usted leyó esta sección el mes pasado, podrá encontrar una cierta similitud entre el mecanismo VCG y la subasta de Vickrey o de segundo precio. De hecho, hay también una subasta VCG que es una generalización de la subasta de segundo precio y que se basa en la misma idea que acabamos de ver: la puja sirve para decidir el adjudicatario, pero la cantidad a pagar no depende de su puja, sino de la del resto de los participantes en la subasta.

Sin embargo, en nuestro ejemplo del ascensor, como en cualquier problema de adquisición de bienes públicos, el mecanismo VCG adolece de un serio inconveniente. Bruno no paga la cantidad que ofreció, d_B, sino una cantidad c_B dada por una fórmula análoga a la de Alicia:

$$c_B = 10.000 - (d_A + d_C + d_D)$$

y así sucede con el resto de los vecinos. Por lo tanto, si se decide hacer la obra, la cantidad total que pagan nuestros cuatro amigos es:

$$C = c_A + c_B + c_C + c_D = 40.000 - 3D$$

en donde D es el total ofertado: $D = d_A + d_B + d_C + d_D$. C es mayor que 10.000 euros sólo si D es menor que 10.000. Pero precisamente la condición para que el

ascensor se construya es que D sea mayor que 10.000 Luego, aunque todos hayan dicho la verdad y la suma de lo declarado supere el costo total, ¡la comunidad no tendrá dinero suficiente para hacer la obra! Algún lector avispado puede haber pensado en la siguiente solución: Alicia da las reglas del mecanismo VCG pero, una vez conocido lo que está dispuesto a pagar cada vecino, las cambia y pide a cada uno lo declarado. Por supuesto, esto es bastante tramposo y la simple sospecha de tal comportamiento rompería el delicado balance sobre el que se basa el mecanismo VCG. Por ello, esta solución es inadmisible.

Una solución más honrada y práctica es completar lo que falta con algún fondo de reserva. Al fin y al cabo, en el caso de que la obra se apruebe, la suma D de lo ofertado no será muy superior a los 10.000 euros. Cada uno de los cuatro vecinos sabe que, si se repartiera el gasto en partes iguales, tendría que pagar 2500 euros. A esta cifra le añadirá o restará alguna cantidad pequeña, dependiendo de su necesidad de ascensor, con lo que la oferta total D estará cercana a los 10.000 y la suma C de las contribuciones será sólo ligeramente inferior al gasto total de 10.000 euros.

A pesar de sus limitaciones, el mecanismo VCG goza de un crédito considerable. Constituye uno de los pilares sobre los que se ha desarrollado todo un campo de la economía: la teoría de *mecanismos de revelación de la demanda*, que se puede aplicar en muy distintas situaciones como decisiones colectivas, subastas, el reparto de pistas en aeropuertos o incluso la regulación del tráfico de datos en Internet.

Las decisiones colectivas en donde se aplican estos mecanismos pueden ser muy variadas. Imaginemos un grupo de amigos que tienen que decidir qué película verán el próximo fin de semana. Supongamos que podemos cuantificar las preferencias de cada uno de ellos en términos monetarios y que éstas son las que se especifican en la siguiente tabla:

Casablanca		*El séptimo sello*	
Alicia	4€	Daniel	2€
Berta	12€	Esther	6€
Carlos	9€	Federico	5€
		Gabriel	7€
Total	**25€**	**Total**	**20€**

Supondremos también que el beneficio de cada individuo es nulo si se va a ver la película que él no desea ver. Aunque la mayoría prefiere *El séptimo sello*, los que quieren *Casablanca* lo desean con más intensidad. Por lo tanto, si quiere maximizar el "beneficio total", el grupo debería ver *Casablanca*. El problema es de nuevo cómo obtener la tabla anterior. Si pedimos a cada miembro del grupo que cuantifique su deseo de ver la película, tenderán a exagerar, puesto que saben que la decisión final depende de la cantidad declarada.

Edward Clarke diseñó un mecanismo para obtener declaraciones sinceras, penalizando las exageraciones con un "impuesto": la *tasa de Clarke*. En primer lugar se calcula la diferencia ΔD entre los dos totales, que en nuestro ejemplo es de 5 euros. Los que están en el bando ganador y han declarado una cantidad superior a esta diferencia deben pagar la tasa, que es la diferencia entre su declaración y ΔD. En nuestro ejemplo, los únicos que tienen que pagar son Berta y Carlos, que abonarán, respectivamente, 7 y 4 euros. Con este "castigo", se puede comprobar que la estrategia óptima de todos los miembros del grupo es decir la verdad. Alicia no tiene interés ninguno en declarar más de los 4 euros con los que valora ir a ver *Casablanca*, puesto que su opción es ya ganadora. Si Berta aumenta su declaración y dice que para ella ver *Casablanca* equivale a un beneficio de 13 euros, acabará pagando la misma tasa. Y lo mismo ocurre si disminuye la cantidad declarada (siempre que no la baje tanto como para que su opción pierda). ¿Qué ocurre con los que están en el bando perdedor? Si, por ejemplo, Gabriel hubiera aumentado su declaración hasta 13 euros, el grupo acabaría asistiendo al pase de *El séptimo sello*, pero ΔD sería 1 euro y Gabriel tendría que pagar 12 euros, con lo que su beneficio neto sería negativo: $7 - 12 = -5$.

En realidad, éste es un caso particular del mecanismo VCG. Si d es mi declaración, D_m indica el total de mi opción y D_c el total de la opción contraria, la tasa de Clarke es:

$$T = d - D_m + D_c = D_c - D^*_m$$

en donde D^*_m representa el total de las declaraciones de mi opción *sin contar la mía*. Como en el mecanismo VCG, la tasa no depende de mi declaración. Si el valor real que tiene mi opción para mí es v, entonces mi beneficio será $U = v - T$ cuando gana mi opción y cero si pierde. Pero mi opción es ganadora si $D_m > D_c$, es decir, si $d > D_c - D^*_m = T$. La situación es idéntica al caso del ascensor: mi beneficio constituye el valor real menos la tasa si la declaración es mayor que la tasa y cero en otro caso.

Uno de los inconvenientes de este mecanismo de toma de decisiones radica en su vulnerabilidad ante coaliciones maliciosas. Federico y Esther podrían ponerse de acuerdo para declarar ambos 20 euros. El total para *El séptimo sello* sería entonces 49 y se convertiría en la opción ganadora. Pero ΔD sería ahora 24, con lo que nadie pagaría tasa alguna. Este aumento concertado de las declaraciones corre su riesgo, pero en general distorsiona el mecanismo de modo que la estrategia "sincera" deja de ser óptima. Por lo tanto, para que el mecanismo resulte eficaz, hay que evitar coaliciones mediante declaraciones secretas y sin posibilidad de negociación previa.

parr@seneca.fis.ucm.es

Geología

Incorporación del tiempo

NICCOLÒ STENONE (1638-1686) ANATOMISTA, GEOLOGO, VESCOVO. Dirigido por Karen Ascani, Hans Kermite y Gunver Skytte. "L'Erma" di Bretschneider; Roma, 2002. THE SEASHELL ON THE MOUNTAINTOP, por Alan Cutler. Dutton; Nueva York, 2003.

BURSTING THE LIMITS OF TIME, por Martin J. S. Rudwick. The University of Chicago Press; Chicago, 2005.

CHARLES DARWIN, GEOLOGIST, por Sandra Herbert. Cornell University Press; Ithaca, 2005.

Tras la revolución copernicana y el descubrimiento de la circulación de la sangre, la ciencia remontó otro hito con la introducción del tiempo en la filosofía natural. Débese esa intuición a Niels Stensen (*Niccolò Stenone* y *The Seashell on the Mountaintop*), sistematizada en teoría por Horace-Bénédict de Saussure (*Bursting the Limits of Time*).

Stensen, latinizado Steno, nació en Copenhague el día de año nuevo de 1638. En su familia abundaban los ministros protestantes, tradición de la que él se desvió para matricularse a los 19 años en la facultad de medicina de la universidad de su ciudad natal. Dos cursos después, cerró la universidad debido al asedio sueco sufrido por Copenhague. Levantado el sitio dos años más tarde, Stensen viajó a Amsterdam para ahondar en anatomía. Al poco de llegar realizó su primer descubrimiento: el conducto excretor de la glándula parótida, el *ductus stenonianus*. De Amsterdam se trasladó a Leiden, entonces magna universidad protestante, donde recalaban muchos estudiantes escandinavos. Desde el punto de vista académico su estancia en Leiden resultó sumamente fecunda. En ese trienio publicó 11 obras, entre ellas una memoria sobre la función de las glándulas exocrinas. Demostró también que el corazón era un músculo. En ese período traba amistad con Baruch Spinoza, refugiado en la cercana Rijnsburg, quien se ganaba el sustento puliendo cristales; ambos comparten una misma afición por la matemática y por el estilo cartesiano de pensar. En otros aspectos, la discrepancia mutua era notable: Spinoza defiende el panteísmo frente al Dios personal de Stensen.

Vuelve a Copenhague, donde sus padres acaban de morir, y publica *Sobre músculos y glándulas*, que dedicó al rey Federico III, confiado en que le valdría para ocupar un puesto en la universidad. No lo consigue, pero hereda una pequeña fortuna de sus progenitores que le permite reemprender el viaje por Europa. En París se le acoge en el círculo de Melchisédec Thévenot, donde coinciden jóvenes naturalistas, matemáticos y humanistas. Disertó allí, a comienzos de 1665, sobre la anatomía del cerebro, texto que constituye una pieza maestra de la historia de la neurología. Es mismo año abandonó París, camino de Montpellier. En la primavera de 1666 cruzó los Alpes. Se detuvo primero en Pisa. Hace gala de su pericia anatómica ante el gran duque Fernando II y su hermano, el príncipe Leopoldo de Medici. Este último, alumno de Galileo, era un amante entusiasta de la ciencia natural que había participado en la fundación, en 1657, de la *Accademia del Cimento*.

Durante su posterior estancia en Florencia goza de la amistad de Francesco Redi, protomédico del gobernante, y del matemático áulico Vincenzo Viviani. Observa el parecido extraordinario entre los dientes del tiburón y las piezas fósiles que se conservan en la Accademia, las "piedras de Malta". La isla, pensó, debió de estar antaño cubierta por el mar. También Toscana, a tenor de las numerosas petrificaciones marinas recuperadas. Pero, ¿cómo pudieron convertirse en piedra? En un opúsculo de 78 páginas (*Prodromus o sobre la inclusión de un cuerpo sólido en otro a través de un proceso natural*) revela que las teorías recibidas en torno al pasado de la Tierra eran erróneas. Para conocer el ayer había que ir levantando, una a una, las capas de la Tierra.

En sus viajes había observado la superposición de estratos, tendiendo un ángulo con el horizonte. De la observación pasó a la reflexión. Esbozó un conjunto de principios que llevan su nombre y trazan los raíles por donde discurrirá la geología. El más famoso es el principio de superposición. Dadas las capas sedimentarias dispuestas una sobre otra, la capa inferior se depositó antes que la superior; cuando se formaba el estrato inferior no existía el superior. Por tanto, los estratos, que registran episodios del pasado, se han formado de un modo secuencial. Suponía que los sedimentos venían transportados por el agua. Un segundo principio, el de horizontalidad, señalaba que todo estrato tuvo un origen horizontal, acorde con el curso del agua; cualquier inclinación o pliegue es consecuencia de acontecimientos posteriores. Con la superposición y la horizontalidad, el tercer principio que se desprende de la lectura del *De solido* declara la continuidad de los estratos, que justificaba por haberse formado por deposición de un mismo curso de agua a ambos lados de lo que hoy es un valle.

En *Bursting the Limits of Time*, obra magistral a la que ha dedicado tres largos decenios de investigación, Rudwick epitomiza en la subida de Horace-Bénédict Saussure (1740-99) al Mont Blanc en agosto de 1787 la convergencia de diversas disciplinas de la historia natural hacia una geología científica. Saussure era



1. Flujo de lava incandescente en mayo de 1771. Dibujo original de Pietro Fabris en la obra de William Hamilton *Campi Phlegraei* (1776).

un “savant”, que había adquirido ya fama con su *Voyage dans les Alpes*, del que se había publicado el año anterior un segundo volumen, con la promesa de la aparición de un tercero. En un arco temporal breve, que cubre las postrimerías del siglo XVIII y comienzos del XIX, se reconstruyó la historia de una Tierra de larga historia con un advenimiento, relativamente reciente, del hombre. Y el autor lo expone con un extenso prenotando vindicativo de la relación complementaria, en ese período, de la fe y la ciencia. No en vano, muchos cultivadores de la historia natural eran pastores protestantes o clérigos católicos. Asentamiento de la geología en esa época significa consolidar la tesis de Stensen y considerar la historia de la Tierra un producto de la propia historia de la naturaleza, un proceso que culminará con su extensión a los seres vivos a través de la teoría darwinista de la evolución.

De particular importancia para la madurez de la geología fueron las escuelas de minería. Tras la devastación de la guerra de los Siete Años se habían establecido tres en Europa central, con la esperanza de reavivar la industria y la economía. En 1765 se creó la Academia municipal de Freiburg; en Berlín se fundó otra, cinco años más tarde. Y una tercera, también 1770, en la ciudad húngara de Schemnitz. En 1773 nació la escuela de san Petersburgo y, en 1777, la española de Almadén. En 1783 se instituía la famosa Ecole des Mines en París. Tales instituciones, que cumplían una misión primordialmente docente, formaban funcionarios. Los alumnos aprendían a descubrir vetas y filones, construir pozos mineros y dominar técnicas de bombeo. Además, las escuelas de minas presentaban un medio de vida para *savants* interesados por la historia de la ciencia de la tierra.

Tal era el caso, por ejemplo, de Abraham Gottlob Werner, profesor de Freiburg, en tanto que su contemporáneo Johann Wolfganf Goethe encontraba su fuente de sustento en la administración de minas. Otros, así Jean-André de Luc, vivían del patronazgo regio. O de su propia fortuna, como el terrateniente James Hutton.

El idioma vehicular de la ciencia, abandonada la preeminencia del latín, empezaba a ser el francés. En ese idioma se relacionaban Saussure y Sir William Hamilton, cuando aquél visitó Nápoles para ver el Vesubio y otras zonas volcánicas de la Campania o cuando le enseñó Chamonix a éste. La Academia de Ciencias de San Petersburgo adoptó el francés por idioma oficial. La propia Regia Sociedad de Londres publicó trabajos en francés, al tiempo que adjuntaba su traducción al inglés en apéndices para comprensión de quienes no entendieran el idioma.

Por lo demás, libros tales como los *Viajes Alpinos* de Saussure y *Campi Phlegraei* de Hamilton trascendían los círculos de los iniciados. Por toda Europa corrían revistas, unas ligadas a corporaciones científicas específicas, como las *Transactions* de la Regia Sociedad o las *Mémoirs* de la Academia Real de Ciencias de París, que sólo contenían artículos previamente debatidos en sus sesiones. Otras publicaciones periódicas eran comerciales y dependían de la voluntad de la casa editora. El modelo más famoso era el *Journal des Savants*, que cada mes informaba de las actividades científicas, y el semanario *Göttingische gelehrte Anzeigen*. No obstante, algunas revistas eran más especializadas, aunque publicaban también generalidades: *Observations sur la Physique*, editada en París. Muy especializada en aspectos mineros era la *Bergmännische Journal*, fundada, en 1778, por la Academia de Minas de Freiburg.

Para los *savants* de la generación de Saussure la ciencia se generaba

en tres sitios fundamentales: el laboratorio, el museo y el campo. Los museos, que habían dejado de ser colecciones de rarezas para convertirse en "inventarios de la naturaleza" donde objetos comunes y excepcionales importaban por igual, ocupaban el lugar de privilegio; los había que constituían gabinetes de aristócratas o burgueses ricos, mientras que otros eran privilegio real. Famoso entre todos, el Museo de Historia Natural de París. Más o menos asociados a esos centros pululaba una masa crítica de artesanos habilidosos que creaban los instrumentos de que se valían los *savants*: barómetros y termómetros, higrómetros y electrómetros, amén de mapas y dibujos de los suelos.

Coexistían en tiempos de Saussure, con frecuentes interrelaciones mutuas, dos enfoques complementarios del mundo natural. Por un lado, la historia natural, que se ocupaba de la descripción y clasificación de fenómenos y entidades (orgánicas e inorgánicas). Por otro, la filosofía natural, que abordaba las relaciones causales y matemáticas entre aquéllos. Modelo de la primera se consideraba el *Systema Naturae* de Linneo; de la segunda, los *Principia* de Newton. Si la historia natural, en cualquiera de sus variantes, se proponía describir y clasificar, cumplía a la filosofía natural descubrir las regularidades y leyes subyacentes.

Ciñéndonos a la historia natural de la Tierra, la mineralogía constituía su núcleo. A imagen de sus homólogas, zoología y botánica, era una ciencia de los especímenes. Para incluirlos debidamente en una ordenación jerárquica, había que empezar por exponer sus caracteres diagnósticos, propiedades físicas y químicas. En una primera partición general había que segregar los minerales simples de los compuestos; a modo de botón de muestra, el granito se encuadraba entre los compuestos pues constaba de cuarzo, feldespato y mica. Aun cuando se sabía ya que los fósiles eran restos de organismos, compartían descripción con minerales y rocas. Algunos empezaron a asociarse con determinadas canteras y estratos.

No tardó, sin embargo, en cuestionarse la primacía de la mineralogía ante la irrupción contemporánea de la geografía física. No bastaba con la labor de gabinete, donde se examinan los especímenes, sino que se exigía trabajar *in situ*. La nueva disciplina estudiaba cadenas montañosas, volcanes, ríos, cuencas de drenaje, valles, fuentes termales, dunas de arena, estuarios. De la litosfera a la atmósfera pasando por la hidrosfera. Saussure se sirve del hidrómetro, toma mediciones atmosféricas de la cumbre del Mont Blanc y describe las tormentas con aparato eléctrico de los Alpes. La geografía física no rivalizaba con la mineralogía, sino que la completaba. Dos elementos en particular reclamaron la atención

Matemática clásica

Espíritu geométrico

TRATADOS, por Arquímedes seguido de **COMENTARIOS** de Eutocio. Vol I. Introducción, Traducción y Notas, Paloma Ortiz García. Biblioteca Clásica Gredos. Editorial Gredos; Madrid, 2005.

Hablar de Arquímedes supone hablar de dos personalidades separadas: la propia de la leyenda y la de una figura histórica conocida de modo fiable sólo por sus contribuciones a la matemática de su época. Plutarco en *Vidas paralelas* inicia la leyenda Arquímedes (–287, –212) en la biografía del general Marcelo, conquistador de Siracusa durante la segunda guerra púnica. A partir de aquí, Arquímedes se convierte en el creador de la física matemática, en el investigador "puro" centrado en la resolución de problemas, pero a la vez en el ingeniero militar o investigador "aplicado" terror de los romanos... Leyenda que se agranda por los altibajos que su figura, como sus trabajos, ha tenido a lo largo de la historia con máxima influencia en los siglos XVI y XVII.

Sus contribuciones a la matemática han sufrido olvidos y recuperaciones. Gracias al interés que despertó en Constantinopla se han conservado sus escritos; interés básicamente de Eutocio, que en el siglo VI pretende poner en claro, hacer comprensibles algunos conocimientos matemáticos. Interés que renace nuevamente en Constantinopla en el siglo IX cuando se compilan, comentan y se tratan de completar los escritos de Arquímedes, se "traducen" al griego bizantino sus trabajos desde el dialecto dorio original. También desde el siglo IX los árabes estudian la obra de Arquímedes. Y gracias a estas dos líneas, bizantina y árabe, se conservan los trabajos arquimedianos que aparecen y desaparecen a lo largo del Medievo, a lo largo de la historia posterior. A esas dos líneas se agrega *El método* contenido en un palimpsesto del siglo X y recubierto en el siglo XIII, que sólo descubriría en 1906 Heiberg en Estambul.

La obra de Arquímedes encuentra varias dificultades aquí, en España, independientes de las intrínsecas a una lectura directa del texto de Arquímedes que no se base en pretendidas reconstrucciones modernas según el esquema del Cálculo diferencial: esos trabajos no han sido traducidos al castellano salvo *El método*. Sólo se puede mencionar una traducción atribuida a Francisco Vera en Argentina carente de valor. Es ahora cuando Paloma Ortiz ha emprendido la no siempre fácil tarea de superar esa laguna y nos ofrece un primer volumen con la traducción de los escritos *Sobre la esfera y el cilindro, Medida del círculo, Sobre los conoides y esferoides*. Esta última obra contiene el teorema que, según la leyenda, Arquímedes pidió que se inscribiera en su tumba. Se agrega una selección de los *Comentarios* de Eutocio, quien recuperó algunos textos arquimedianos en el siglo VI. Una selección que enriquece el valor de este libro.

En esta traducción se manifiesta con nitidez el espíritu geométrico de la matemática griega, apoyado en dos pilares: la figura o ilustración, el empleo de la teoría de las proporciones. Para entender rectamente los escritos de Arquímedes, y de los matemáticos griegos en general, es imprescindible el dominio de esta última teoría, muy alejada del tipo de razonamiento más algebraico que manejamos hoy. Alejamiento que exige por parte del lector actual un esfuerzo para captar y comprender los pasos del razonamiento matemático en su sentido griego primigenio, que se arropa en lo exterior, en lo formal, en las demostraciones por reducción al absurdo, en el esquema axiomático "euclídeo". A pesar de sus dificultades, nunca viene mal traducir una obra de indudable valor histórico, ponerla en castellano al alcance de todo el que ame la historia del pensamiento y, en particular, del pensamiento matemático. Con la esperanza de que no se demore la edición de los restantes trabajos de Arquímedes. —JAVIER DE LORENZO

de los geógrafos físicos: volcanes y terremotos. Si abundan los escritos sobre las erupciones del Vesubio, no menor inquietud científica mostraron los *savants* ante el terremoto sufrido por Calabria en 1783. Al lugar acudieron muchos para tomar apuntes de los efectos producidos. Anotaciones que se reflejaban en la cartografía. Jean-Etienne Guettard publicó en 1784 un mapa de Francia en el que distinguía tres tipos de suelos: arenoso, margoso y pizarroso.

Todavía floreció una tercera ciencia descriptiva de la tierra. Relacionada con la geografía física, hizo de la dimensión de profundidad su foco de investigación. Se difundió bajo distintas denominaciones: geognosia, geografía mineralógica, geografía subterránea o anatomía de la Tierra. Tales apelativos reflejan su vinculación directa con el mundo de la minería, que le aportaba los datos empíricos sobre la distribución vertical de las rocas. La composición de los estratos de pozos o cortes se ilustraba en los primeros mapas coloreados. La geognosia acotaba su dominio de estudio en unidades topográficas, cuya descripción incluía no sólo un registro de las clases de rocas que la componían, sino también sus relaciones estructurales y los tipos de composición topográfica. La clasificación en geognosia era también jerárquica. Tras establecer la "especie" como taxón natural fundamental de la masa rocosa, que se suponía de validez universal, la taxonomía podía extenderse hacia arriba para abarcar agrupaciones más comprehensivas y fundamentales o hacia abajo, hacia variedades locales. A la distinción fundamental entre masas rocosas primarias y secundarias, se añadió la partición entre masas rocosas aluviales y masas rocosas volcánicas.

Competía a la filosofía natural, a la "física natural" desentrañar las regularidades o leyes naturales que subyacían bajo fenómenos, procesos y entidades terrestres, con el fin de determinar las causas físicas. ¿A qué se debía, por ejemplo, que determinadas rocas primarias presentaran una estratificación tal, que parecían secundarias? ¿Se habían formado capa a capa como precipitados químicos? Piedra de contradicción resultó ser, en particular, el granito, que a menudo se manifiesta como una estructura



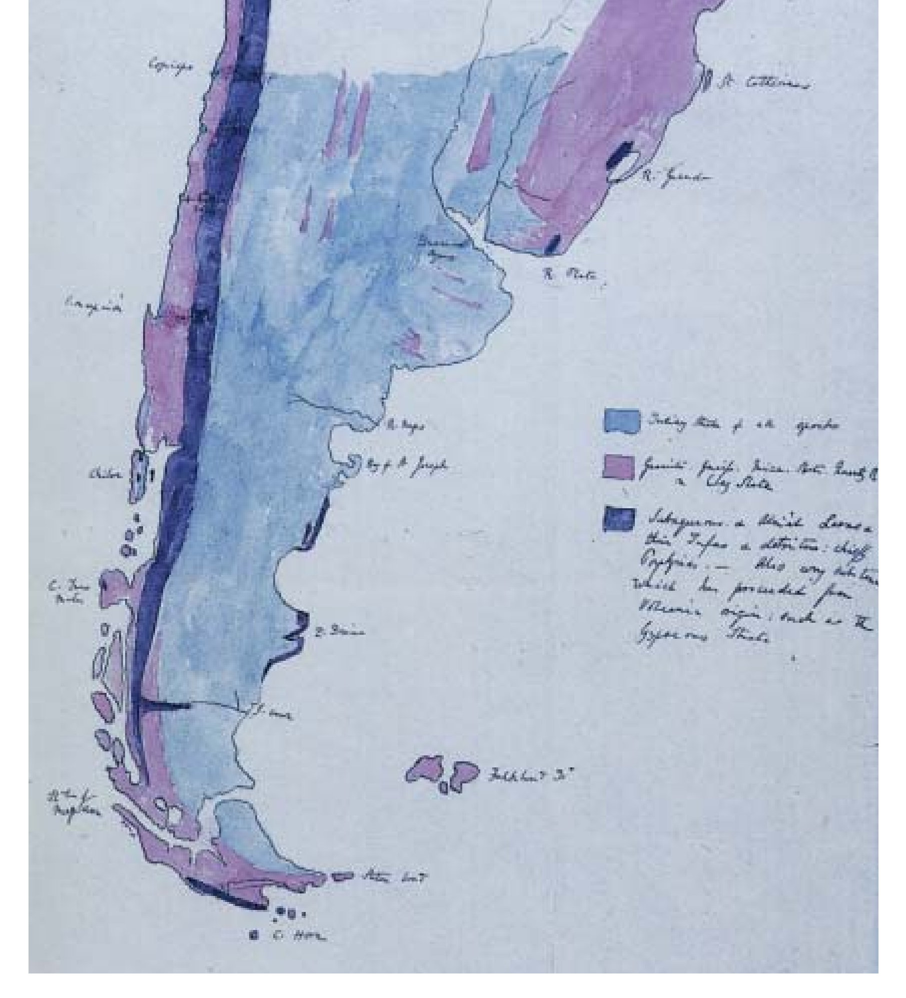
2. Especímenes diversos. Museo Sedgwick. Universidad de Cambridge.

en capas, lo mismo que el mármol; en busca de una razón se invocaba el ejemplo de las lechadas de hornos y fundiciones. ¿Se forjaron en las entrañas ardientes del interior de la Tierra? Aceptado, por otro lado, el origen orgánico y en muchos casos marino de los fósiles, su actual ubicación se atribuía o bien a un episodio violento transitorio o a una situación distinta de mares y continentes en el pasado. Urgía, a su vez, una explicación de esas revoluciones que convertían mares en continentes y continentes en mares.

Muy ingeniosa se mostró la física natural a la hora de dar cuenta de accidentes menores. Así la formación de valles, atribuidos en unos casos a la erosión y en otros a fenómenos violentos de corrimientos de tierras o fallas. Para justificar el extraño caso de los bloques errantes, peñascos sin ninguna afinidad mineralógica con las rocas circundantes, muchos *savants*, incluido Saussure, los relacionaban con un episodio de excepcional violencia y repentino, una suerte de megatsunami. En la búsqueda de las causas de volcanes y terremotos, no se le escapó a John Michell su mutua asociación frecuente en el tiempo y en el espacio. Al estudio de los volcanes solía ir engarzada la cuestión del origen del basalto, objeto de controversia entre plutonistas y neptunistas. En las corrientes de lava subterránea



3. Especímenes de rocas de la isla James. Museo Sedgwick. Universidad de Cambridge.



4. Dibujo en color de América del Sur realizado por Charles Darwin.

fiaban los vulcanistas el origen de las vetas minerales que cruzaban las rocas primarias, en tanto que para el neptunista Werner se trataba de hendiduras en la corteza terrestre, rellenadas por materiales precipitados. Todos los procesos causales esgrimidos compartían un carácter común, la dimensión temporal.

Más allá de las explicaciones causales de la elevación de las montañas, el origen de los volcanes, la consolidación de las rocas o la situación de los fósiles, los *savants* soñaban con alcanzar una teoría cabal, un sistema. Debía éste ser capaz de explicar los principales rasgos distintivos de la Tierra, concatenar las causas naturales con sus efectos, atenerse a entidades físicas conocidas y procesos observables y ofrecer una explicación plausible del desarrollo pasado, presente y futuro de la Tierra.

Darwin fue uno de los que soñaron con crear su propia teoría sistemática de la ciencia de la Tierra (*Charles Darwin, Geologist*), basada en los movimientos de la corteza terrestre. Es más, la geología desempeñó un papel determinante en la gestación del *Origen de las especies*, fruto de un fecundo viaje de circunnavegación a bordo del *H.M.S. Beagle* (1831-1836), para cuya realización se inspiró en Werner vía Alexander Humboldt. Su sueño lo llevó a la realidad sólo en parte, con su teoría de la estructura y distribución de los arrecifes coralinos.

Cuando en 1859 la Sociedad Geológica de Londres otorga a Darwin su máximo galardón, la medalla Wollaston, por sus contribuciones a las ciencias de la Tierra, Charles Lyell reseña sus trabajos sobre el crecimiento de los arrecifes coralinos, sobre las pruebas de la elevación moderna de Chile y la elevación de los Andes, la distribución de los bloques erráticos en Bretaña y los invertebrados fósiles. Mucho debía Darwin a Lyell en geología. De hecho, apenas si se cuestiona que el darwinista *Sobre el origen de las especies* constituya la prolongación de los *Principles of Geology* (1830-1833) de Lyell. Pero si demostrado está el influjo del padre del uniformismo, no menos peso ejercieron en la formación del joven Darwin William Buckland, Sedgwick, Henslow, Humboldt y otros. De Robert Jameson recibió Darwin la enseñanza de los fundamentos de la mineralogía y ciertos rudimentos de geología de campo, de acuerdo con un método comprehensivo y sistemático. Jameson indicaba, de cada espécimen, su figura cristalina, grado de dureza (escala de 1 a 10, del talco al diamante) y gravedad específica; añadía sus circunstancias geognósticas y geográficas. A Segdwick debía Darwin sus primeros escarceos de geología de campo, pertrechado de un clinómetro que le había proyectado Henslow.

La obra geológica del padre de la teoría de la evolución está jalonada por la trilogía de su viaje: *The Structure and Distribution of Coral Reefs* (1842), *Volcanic Islands* (1844) y *Geological Observations on South America* (1846). A lo largo de 1846 Darwin publicó 19 artículos o noticias sobre asuntos geológicos, muchos de los cuales los presentó ante la Sociedad Geológica de Londres. Desde ese momento empezó a distanciarse de los temas geológicos. Al explicar la elevación continental de Sudamérica, declaraba que las llanuras meridionales se habían ido alzando a través de sucesivas elevaciones propiciadas por fuerzas que actuaban en una zona extensa. También se interesó por la pendiente del suelo marino adyacente a la línea de costa de la Patagonia. En su opinión, la elevación de las llanuras requería la expansión gradual de la masa central de la Tierra que operaba por intervalos sobre la corteza externa.

En su teoría de la estructura y distribución de los arrecifes de coral, resaltó la función de la subsidencia en la creación de las islas de coral, y de forma llamativa en los atolones. Los naturalistas franceses Jean René Constant Quoy y Joseph Paul Gaimard habían observado que los organismos que creaban arrecifes coralinos operaban a una profundidad relativamente somera. Darwin añadía que la formación del coral coincidía con el hundimiento del suelo. En su posterior desarrollo consideró que la formación de los arrecifes coralinos era un equilibrio entre la elevación de América del Sur y la subsidencia de la cuenca del Pacífico.

—Luis Alonso



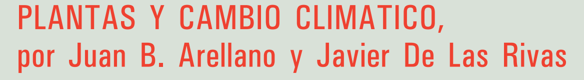
EL GENOMA AL ALCANCE DE TODOS, por George M. Church

En menos de una década llegarán nuevas técnicas de secuenciación, más rápidas y baratas, que facilitarán el acceso a la información génica.



EL GLOBO COMETARIO DE VINCENZO CORONELLI, por Wilhelm Seggewiss

Los globos celestes de Tréveris construidos por Vincenzo Coronelli seducen por sus dimensiones, pero también por su síntesis de esplendor barroco y sus elevadas pretensiones científicas.



PLANTAS Y CAMBIO CLIMATICO, por Juan B. Arellano y Javier De Las Rivas

Los gases de invernadero constituyen una amenaza contra la integridad de la biosfera. Para hacer frente a ese reto, se cuenta con dos complejos enzimáticos de la fotosíntesis.



EL CEREBRO MATERNO, por Craig Howard Kinsley y Kelly G. Lambert

El embarazo y la maternidad cambian la estructura cerebral de las hembras de mamífero. Ello hace que las madres sean más diligentes con sus pequeños y les presten mejores cuidados.



EL ORIGEN DE LAS ENANAS MARRONES, por Subhanjoy Mohanty y Ray Jayawardhana

Son una astilla para las teorías de formación de estrellas y planetas; aun así, las enanas marrones pueden ayudar a resolverlas.